AL

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
1 August 2002 (01.08.2002)

PCT

(10) **International Publication Number**
WO 02/059369 A2

(51) **International Patent Classification**[7]: **C12Q 1/68**, A61K 31/70, C12N 15/64, 15/11, G01N 33/53

(21) **International Application Number:** PCT/US01/47962

(22) **International Filing Date:** 1 November 2001 (01.11.2001)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
60/245,031 1 November 2000 (01.11.2000) US

(71) **Applicant: UNIVERSITY OF MEDICINE AND DENTISTRY NEW JERSEY MEDICAL SCHOOL** [—/US]; Department of Medicine, P.O. Box 2688, New Brunswick, NJ 08903-2688 (US).

(72) **Inventor: LOBEL, Peter**; 411 Benner Street, Highland Park, NJ 08904 (US).

(74) **Agent: WISE, Michael, J., Esq**; Lyon & Lyon LLP, Suite 4700, 633 West Fifth Street, Los Angeles, CA 90071 (US).

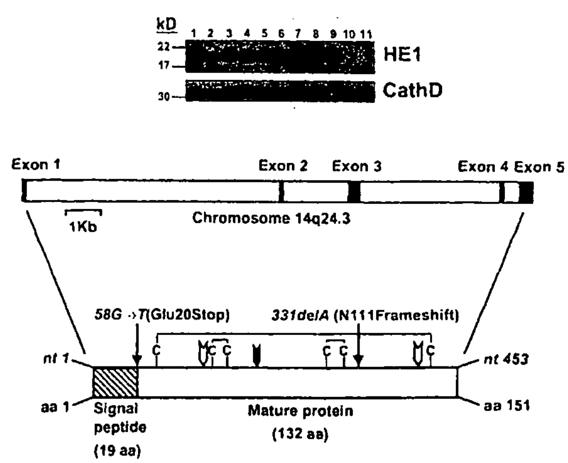
(81) **Designated States** *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
– *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** COMPOSITIONS AND METHODS FOR THE TREATMENT OF DISEASES RELATED TO FAULTY CHOLESTEROL REGULATION



(57) **Abstract:** Niemann-Pick type C2 (NP-C2) disease is a fatal lipid storage disorder characterized by massive lysosomal accumulation of cholesterol. The present invention identifies HE1 as the gene responsible for NP-C2. Treatment of NP-C2 fibroblasts with an exogenous HE1 genetic element ameliorated the cholesterol accumulation phenotype. HE1 functions in intracellular cholesterol transport. The present invention provides therapeutic compositions consisting of HE1 polynucleotide and polypeptide sequences, as well as an expression system for expressing HE1 in target cells. These therapeutic compositions can be used to target diseases involving faulty cholesterol transport and regulation, including NP-C2, atherosclerosis, Alzheimer's, diabetes, and cardiovascular disease. In addition, the present invention provides methods of diagnosing both NP-C2 and the ability of a subject to genetically transmit the disease by detecting mutations in the HE1 gene sequence.

WO 02/059369 A2

COMPOSITIONS AND METHODS FOR THE TREATMENT OF DISEASES RELATED TO FAULTY CHOLESTEROL REGULATION

BACKGROUND OF THE INVENTION

[0001] The present utility patent application claims priority to provisional patent application U.S. Ser. No. 60/245,031 (Lobel), filed November 1, 2000, the disclosure of which is incorporated by reference in its entirety herein.

GOVERNMENT INTEREST

[0002] This invention was made with government support by the following Public Health Service grants: DK54317, NS37918, and DK45992 from the National Institute of Health. The government may own certain rights in the present invention.

FIELD OF THE INVENTION

[0003] The present invention relates to the field of molecular biology, cholesterol storage disease, cholesterol regulation, and cardiovascular disease. In particular, this invention provides the genetic basis for Niemann-Pick C2 disease and sets forth novel compositions and methods with which to treat this disease as well as other diseases associated with faulty cholesterol regulation, such as cardiovascular disease, atherosclerosis, Alzheimer's disease and diabetes. This is due to the fact that the gene responsible for Niemann-Pick C2 disease, HE1, plays a vital role in cellular cholesterol trafficking.

BACKGROUND

[0004] Niemann-Pick disease is a general term for a group of diseases that affect metabolism. The three most common recognized forms of the disease are types A, B, and

C (NP-A, NP-B, and NP-C). NP-A and NP-B are autosomal recessive disorders linked to a deficiency in the enzyme sphingomyelinase (ASM) (1). In a subject with NP-A or NP-B, ASM does not properly metabolize sphingomyelin, leading to an accumulation of this lipid within the cell (2). Although both NP-A and NP-B result from deficiencies in the same enzyme, they exhibit drastically different clinical results. Patients with NP-A exhibit severe neurological effects, generally leading to death by 2 to 3 years of age, while patients with NP-B generally have little or no neurological involvement, and may survive into late childhood or adulthood (2).

[0005] Niemann-Pick disease type C (NP-C) is an autosomal recessive lipid storage disorder characterized by progressive deterioration of the central nervous system, visceral symptoms, and premature death (3). There are two forms of NP-C disease, NP-C1 and NP-C2. NP-C1 is substantially more common than NP-C2, with a rate of occurrence estimated at 1 in 10^5 live births, as opposed to an occurrence of only 8 known cases of NP-C2 worldwide. At the cellular level, the most prominent feature of the NP-C lesion is lysosomal sequestration of LDL-derived cholesterol, which results in downstream effects on cholesterol homeostasis (4, 5). In addition, NP-C cells demonstrate a deficiency in both H and L ferritins, suggesting that they prevent utilization of iron for the synthesis of cytosolic ferritin (6).

[0006] In a healthy subject, LDL particles transport dietary cholesterol to fibroblasts, where they enter the cell via LDL receptor-mediated endocytosis. After entering the cell, vesicles containing the LDL fuse with lysosomes. The LDL particle is degraded, and free cholesterol is released into the cytosol (7). Little is known about how free cholesterol is

transported from the lysosome to the endoplasmic reticulum, plasma membrane, and other cellular sites (8).

[0007] In patients with NP-C, cholesterol transport is faulty, which results in accumulation of LDL-derived cholesterol within the lysosome. Because of this, NP-C cells exhibit massive cholesterol storage as demonstrated by filipin staining (9, 10). In addition, NP-C cells display an increase in endogenous cholesterol synthesis, leading to further cholesterol accumulation. In a healthy cell, cholesterol released from the lysosome enters a feedback loop whereby it downregulates endogenous cholesterol synthesis by inhibiting the activity of HMG-CoA. In NP-C cells, this feedback cycle is defective because free cholesterol is not released by the lysosome. The result is increased cholesterol synthesis, leading to a further accumulation of cholesterol within the cell.

[0008] Somatic cell hybridization experiments using skin fibroblast cultures from unrelated NP-C patients demonstrated the existence of a major complementation group comprising ~95% of cases, designated NP-C1, and a minor complementation group, designated NP-C2 (9, 10, 11). The NP-C1 gene has been identified and mapped to chromosome 18q11 (12). The defect responsible for NP-C2 has been excluded from this region by linkage analysis (9). NP-C patients from both complementation groups demonstrate similar clinical and biochemical phenotypes, suggesting that the genes responsible for each may interact or function sequentially in a common metabolic pathway.

[0009] The gene responsible for NP-C1 disease has been cloned and characterized (12). NPC1 encodes an integral transmembrane protein consisting of 1278 amino acids, with a

lysosomal targeting motif and a putative sterol-sensing domain (12). Despite extensive investigation, the precise function of NPC1 remains elusive. NPC1 resembles a family of bacterial permeases that transport various substances, including fatty acids, through the bacterial cell membrane (13). There is evidence that NPC1 has a permease-like pump function, but the implication for cholesterol transport is unclear (13).

[0010] NP-C2 is extremely rare, with only eight total cases having been reported worldwide (14). The cases reported thus far arose in France, Algeria, Italy, Germany, the Czech republic, and Turkey (14). One of these cases exhibited a milder phenotype, with the patient displaying juvenile onset of neurological symptoms and prolonged survival (14). In the other seven cases, patients suffered a severe and rapid disease course, with all patients dying between 6 months and 4 years of age (14). Six of these seven cases displayed pronounced lung involvement, with patients dying of respiratory failure (14). Two of these patients also developed severe neurological disease during infancy (14). There is currently no available therapy for the treatment of NP-C2. The devastating nature of the disease, along with its apparent link to the crucial process of intracellular cholesterol transport, make the need for therapeutic development extremely important.

[0011] The present invention identifies HE1 as the gene responsible for NP-C2 disease. HE1 was originally cloned by differential screening of a human epididymal cDNA library, and found to be highly expressed in all parts of the human epididymis (15, 16). The complete sequence of HE1 is available (AC005479). Characterization of the full-length HE1 gene revealed that it is a single copy gene encoding a 151 amino acid glycoprotein of 25-27 kDa (17).. The protein contains a 19 amino acid sequence that is highly conserved

among mammalian species and which represents a major secretory component of epididymal fluid (18, 19, 20, 21). HE1 is present in numerous cDNA and SAGE libraries (see UniGene cluster Hs.119529, www. ncbi.nlm.nih.gov/UniGene). In addition, a bovine homolog is present in milk (22) and bovine and murine messages are detected in several tissues (22, 23). These observations suggest that, in addition to its speculated role in sperm maturation (19, 24), HE1 may have a more global function. Northern blot analysis has revealed HE1 mRNA in all human tissues examined, with highest levels in the testis, kidney, and liver, and lowest in lung and muscle (25). It has recently been reported that the porcine homolog of HE1 specifically binds cholesterol (19).

SUMMARY OF THE INVENTION

[0012] In summary, Niemann-Pick disease type C2 (NP-C2) is a fatal hereditary disorder of unknown etiology characterized by defective egress of cholesterol from lysosomes. The present inventor has discovered that HE1, a previously described cholesterol binding protein isolated from epididymal fluid, is a ubiquitously expressed lysosomal protein. This finding is consistent with the presence of HE1 in epididymal fluid, because epididymal fluid has an acidic milieu and is an abundant source of several other lysosomal proteins (18, 26, 27, 28). Further, the present inventor has discovered that a defect in HE1 protein expression is responsible for the accumulation of LDL-derived cholesterol in NP-C2 patients. Based on the work of the present inventor, it was determined that HE1 is undetectable in fibroblasts from NP-C2 patients, but present at normal levels in control patients. In addition, HE1 levels are elevated in NP-C1 patients. These observations led the present inventor to develop a treatment for subjects affected by NP-C2, and as

described herein below, the treatment of NP-C2 (HE1 deficient) fibroblasts with exogenous recombinant HE1 protein ameliorates accumulation of LDL-derived cholesterol.

[0013] The exact role of HE1 in cholesterol transport has still not been fully characterized. The present inventor suggests that one role for HE1 can be as a carrier that prevents unregulated intercalation of cholesterol into lipid bilayers, selectively transferring sterols to transmembrane proteins such as NPC1. By facilitating this transfer, HE1 ensures that free LDL-derived cholesterol can escape from the lysosomal compartment. This mechanism explains why NP-C2 cells, in which HE1 is inactive, exhibit large-scale cholesterol accumulation.

[0014] Hence, although the role of normal HE1 in cholesterol transport is still not completely clear, the present discovery identifies defective HE1 expression as the cause of NP-C2, and thus implicates this defective expression in faulty cholesterol transport. This discovery of the molecular basis for NP-C2 facilitates genetic, biochemical, and physical studies to further elucidate the process and mechanism of lysosomal cholesterol transport.

[0015] In addition, discovery of the role of HE1 in NP-C2 provides an immediate therapeutic and diagnostic weapon against both NP-C2 and other diseases involving faulty cholesterol regulation. Potential targets for HE1-based therapies include atherosclerosis, Alzheimer's disease, diabetes, and cardiovascular disease. Each of these diseases is linked to defective cholesterol regulation, and an increase in patient cholesterol levels. In Alzheimer's disease, high cholesterol levels appear to increase formation of ß-amyloid protein, which has been implicated as the primary cause of brain degeneration (29).

Research with both cultured cells and animals has shown that statins (drugs designed to lower cholesterol levels) appear to lower the prevalence of Alzheimer's disease. However, statins are not an ideal therapy for decreasing Alzheimer's, due to their sometimes dangerous side effects, which include liver damage and rhabdomyolosis (29). In atherosclerosis, high blood cholesterol levels lead to the formation of plaques on artery walls, which can decrease blood flow and increase the risk of stroke. The incidence of atherosclerosis is markedly increased in patients suffering from diabetes. Cardiovascular disease is also marked by a substantial increase in cholesterol levels. The connection between HE1 and cholesterol regulation disclosed herein suggests that HE1 therapy may be useful in combating these and other diseases in addition to NP-C2 disease.

[0016] The discovery of HE1 as the second gene of Niemann-Pick type C disease has led to the compositions and methods of the present invention. The present invention provides pharmaceutical compositions containing the HE1 polynucleotide polypeptide and antisense polynucleotide sequence, as well as vectors for their delivery to target cells and expression systems for producing the HE1 polypeptide. In addition, the present invention provides methods for both diagnosing and treating NP-C, as well as a method for treating other diseases in which cholesterol regulation is faulty.

[0017] According to one aspect of this invention, a pharmaceutical composition consisting of the polynucleotide encoding HE1 is provided. Preferably, the polynucleotide sequence is substantially similar, if not identical, to SEQ. ID. NO: 1, or is a sequence with at least 70% identity to SEQ. ID. NO: 1 as its antisense sequence. The invention further includes a pharmaceutical composition consisting of the polypeptide sequence of HE1. Preferably,

this polypeptide sequence is identical to the sequence of SEQ. ID. NO: 2, or is a sequence substantially similar to SEQ. ID. NO: 2. In addition, the present invention includes a pharmaceutical composition consisting of a vector containing the HE1 polynucleotide sequence that encodes the HE1 polypeptide sequence that is substantially similar, if not identical, to SEQ. ID. NO: 2.

[0018] Another aspect of the invention is an expression system for producing the HE1 polypeptide of SEQ. ID. NO: 2. This expression system consists of an HE1 therapeutic element that in one embodiment contains an expression cassette, which includes a HE1 genetic element (e.g., a HE1 DNA, cDNA, RNA, antisense polynucleotide sequence, polypeptide or protein) along with one or more additional elements, and a delivery vehicle such as a vector. Preferably, the delivery vehicle can be a plasmid, cosmid, bacteriophage, or virus, and includes all the appropriate, additional regulatory elements.

[0019] The present invention may be implicated not only in NP-C2, but also in other diseases involving faulty cholesterol regulation. Hence, the invention relates to compositions and methods of treating both NP-C2 and other diseases linked to faulty cholesterol regulation, such as atherosclerosis. Treatment of these conditions is preferably achieved by administering an HE1 genetic element such as the polynucleotide sequence of SEQ ID NO: 1 or the polypeptide or protein of SEQ. ID. NO: 2. Administration of the HE1 genetic element can occur either by introducing purified polynucleotide or protein directly or more preferably by introduction of an expression system capable of producing the HE1 polypeptide within the subject's cells. Preferably, this expression system consists of an HE1 therapeutic element that includes an expression cassette containing an HE1

genetic element along with one or more additional elements, and a delivery vehicle. In one preferred embodiment, this delivery vehicle may be a vector such as a plasmid, cosmid, or virus, which may contain one or more additional elements.

[0020] Further, the invention provides for a method of diagnosing NP-C2 within a subject by detecting a mutation in the HE1 gene of sequence SEQ. ID. NO: 1. Discovery of such a mutation, be it a deletion, substitution, splice mutation, or insertion, indicates the presence of NP-C2 in said subject. Preferably, the method of detecting the mutation consists of either amplifying the HE-1 gene sequence and performing sequence analysis or hybridizing the HE1 gene sequence with a labeled nucleic acid probe corresponding to the wild type HE1 nucleotide sequence of SEQ. ID. NO: 1. In addition to providing a means of diagnosing NP-C2 in a patient, the invention also provides a method of applying the above sequence analysis to detect the potential of a subject to genetically transmit NP-C2. This embodiment of the invention is based on the idea that a subject with a single allele mutation in the HE1 gene may be able to transmit the disease without displaying any symptoms of NP-C2. The present invention also features a method of diagnosing NP-C1 in a subject by detecting elevated expression levels of the HE1 gene. This increased expression is preferably detected by analysis of either HE1 mRNA levels or HE1 protein levels within a subject.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021] The invention is best understood from the following detailed description when read in connection with the accompanying drawings, in which:

[0022] FIG. 1. Analysis of human brain mannose 6-phosphorylated glycoproteins. **A.** Coomassie blue-stained polyvinylidine difluoride membrane. **B.** Nitrocellulose membrane probed with antibodies to HE1 and visualized by chemiluminescence. **C.** Nitrocellulose membrane probed with 2nM ^{125}I-labeled soluble cation-independent MPR. **D.** HE1 expression in various human tissues.

[0023] FIG. 2. Subcellular distribution of HE1 homolog in rat liver. **A.** HE1 distribution in differential centrifugation fractions. **B.** HE1 distribution in sucrose density gradient fractions from control rats. **C.** HE1 distribution in sucrose density gradient fractions from rats treated the nonionic detergent Triton WR1339.

[0024] FIG. 3. **A.** HE1 and cathepsin D protein levels in control and mutant fibroblasts. Lane 1-3, fibroblasts from unaffected control subjects; Lanes 4-5, fibroblasts from sea blue histiocyte disease subjects; Lanes 6-9, fibroblasts from NP-C1 disease subjects; Lanes 10-11, fibroblasts from NP-C2 disease subjects. **B.** Schematic of the HE1 gene and protein.

[0025] FIG. 4. Correction of cholesterol accumulation in NP-C2 fibroblasts. **A.** Cholesterol accumulation in the absence of supplement. **B.** Cholesterol accumulation in presence of 0.3% conditioned medium from a CHO cell line producing recombinant human HE1. **C.** Cholesterol accumulation in presence of 0.3% conditioned medium from untransfected CHO cells. **D.** Western blot comparison of HE1 expression HE1-transfected CHO cells (lane 2) and untransfected CHO cells (lane 1). **E.** Fluorescence measurements of cells from **A**, **B**, and **C**.

DETAILED DESCRIPTION

Introduction

[0026] Applicants have identified HE1 as the gene responsible for Niemann-Pick type C2 disease. Niemann-Pick disease type C (NP-C) is an autosomal recessive lipid storage disorder characterized by progressive deterioration of the central nervous system, visceral symptoms, and premature death (3). There are two forms of NP-C disease, NP-C1 and NP-C2. NP-C1 is substantially more common than NP-C2, with a rate of occurrence estimated at 1 in 10^5 live births, as opposed to an occurrence of only 8 known cases of NP-C2 worldwide. At the cellular level, the most prominent feature of the NP-C lesion is lysosomal sequestration of low density lipoprotein (LDL)-derived cholesterol, resulting in downstream effects on cholesterol homeostasis (4, 30). Somatic cell hybridization experiments with skin fibroblast cultures from unrelated NP-C patients demonstrated the existence of a major complementation group comprising ~95% of cases, designated NP-C1, and a minor complementation group, designated NP-C2 (9, 10, 11).

[0027] Discovery of the molecular defect responsible for producing NP-C2 arose from a study directed at characterizing the lysosome proteome. The approach that led to the discovery was based on the fact that many soluble lysosomal proteins acquire a posttranslational modification that distinguishes them from most other types of proteins, the mannose 6-phosphate (Man6-P) marker. This modification is recognized by Man6-P receptors (MPRs), which divert newly synthesized lysosomal enzymes from the secretory pathway to the endolysosomal system (33). Purified MPR derivatives typically bind phosphorylated lysosomal proteins with subnanomolar affinity and can be used to detect and purify Man6-P glycoproteins (31, 34).

[0028] A two-dimensional gel map of MPR affinity-purified proteins from human brain contained a group of proteins sharing the same NH_2-terminal sequence, as determined by Edman degradation (longest sequence, EPVQFKDXGSVDGVIK), which are likely to represent differentially glycosylated isoforms of the same protein. This sequence perfectly matched the processed NH_2-terminus of HE1 (16, 17), a 151-amino acid glycoprotein containing a 19-amino acid signal that, along with homologs from numerous mammalian species, represents a major secretory component of epididymal fluid (18, 19, 20, 21). Western blotting with polyclonal antibodies against recombinant HE1 confirmed the identity of the proteins, and probing with radiolabeled MPR verified that HE1 contained the Man6-P modification.

[0029] Analysis of HE1 mRNA by Northern blotting revealed a single transcript of 0.9 kb in all tissues examined, with highest levels in testis, kidney, and liver and lowest levels in lung and muscle. This wide distribution is consistent with the presence of HE1 sequences in numerous cDNA and SAGE libraries [see UniGene cluster Hs.119529]. In addition, a bovine HE1 homolog is present in milk (22), and bovine and murine HE1 mRNAs have been detected in several tissues (22, 23). These observations suggest that, in addition to its postulated specialized role in sperm maturation (19, 24), HE1 may have a more global function.

Definitions

[0030] Various terms relating to the biological molecules of the present invention are used throughout the specification and claims.

[0031] "HE1" refers generally to an HE1 polypeptide that is ubiquitously expressed in the lysosome. This HE1 polypeptide is absent in fibroblasts taken from patients with NP-C2 disease, but present at normal levels in cells taken from healthy individuals. HE1 is a cholesterol binding protein that appears to play a central role in intracellular cholesterol transport and regulation. "HE1 activity or HE1 polypeptide activity" or "biological activity of the HE1 or HE1 polypeptide" refers to the metabolic or physiologic function of said HE1 including similar activities or improved activities or these activities with decreased undesirable side effects.

[0032] "HE1 gene" refers to a polynucleotide as defined above in accordance with the present invention, which encodes an HE1 polypeptide.

[0033] An "HE1 therapeutic" refers to a therapeutically effective amount of an HE1 related genetic sequence such as, but not limited to polynucleotide, polynucleotide antisense sequence, and HE1 peptide, protein or protein fragment.

[0034] "Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

[0035] "Polynucleotide" generally refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA,

DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, RNA that is a mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA.

[0036] The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications have been made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

[0037] "Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides, or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the gene-encoded amino acids. "Polypeptides" include amino acid sequences modified whether by natural processes, such as posttranslational processing, or by chemical modification techniques that are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well

as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains, and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods.

[0038] Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racernization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. See, for instance, *Proteins - Structure And Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993* and *Wold, F., "Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in "Posttranslational Covalent Modification Of Proteins", B, C. Johnson, Ed., Academic Press, New York, 1983*; *Seifter et al., "Analysis for protein modifications and nonprotein*

cofactors", Meth Enzymol (1990) 182:626-646 and *Rattan et al., "Protein Synthesis: Posttranslational Modifications and Aging", Ann AIYAcad Sci (1992) 663:48-62.*

[0039] "Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical.

[0040] The term "substantially similar" refers to amino acid sequences having sequence variations that do not materially affect the nature of the protein (i.e. the structure, stability characteristics, substrate specificity, and/or biological activity of the protein). With regards to amino acids, "substantially similar" refers generally to conservative substitutions and/or variations in regions of the polypeptide not involved in determination of structure or function.

[0041] The term "percent identity" or "percent identical" with regards to nucleic acid molecules refers to the percentage of the nucleotides of the subject nucleic acid sequence that have been matched to identical nucleotides by a sequence analysis program.

[0042] "Identity" and "similarity" can be readily calculated by known methods. Nucleic acid sequences and amino acid sequences can be compared using computer programs that align the similar sequences of the nucleic or amino acids thus define the differences. In preferred methodologies, the BLAST programs (NCBI) and parameters used therein are employed, and the DNAstar system (Madison, WI) is used to align sequence fragments of genomic DNA sequences. However, equivalent alignments and similarity/identity assessments can be obtained through the use of any standard alignment software. For instance, the GCG Wisconsin Package version 9.1, available from the Genetics Computer Group in Madison, WI, and the default parameters used (gap creation penalty=12, gap extension penalty=4) by that program may also be used to compare sequence identity and similarity.

[0043] With respect to single-stranded nucleic acid molecules, the term "specifically hybridizing" refers to the association between two single-stranded nucleic acid molecules of sufficiently complementary sequence to permit such hybridization under pre-determined conditions generally used in the art (sometime termed "substantially complementary"). In particular, the term refers to hybridization of an oligonucleotide with a substantially complementary sequence contained within a single-stranded DNA or RNA molecule, to the substantial exclusion of hybridization of the oligonucleotide with single-stranded nucleic acids of non-complementary sequence.

[0044] With respect to oligonucleotide constructs, but not limited thereto, the term "specifically hybridizing" refers to the association between two single-stranded nucleotide molecules of sufficiently complementary sequence to permit such hybridization under pre-

determined conditions generally used in the art (sometimes termed "substantially complementary"). In particular, the term refers to hybridization of an oligonucleotide construct with a substantially complementary sequence contained within a single-stranded DNA or RNA molecule of the invention, to the substantial exclusion of hybridization of the oligonucleotide with single-stranded nucleic acids of non-complementary sequence.

[0045] The term "substantially pure" refers to a "preparation comprising at least 50-60% by weight the compound of interest (e.g., nucleic acid, oligonucleotide, protein, etc.). More preferably, the preparation comprises at least 75% by weight, and most preferably 90-99% by weight, the compound of interest. Purity is measured by methods appropriate to the compound of interest (e.g. chromatographic methods, agarose or polyacrylamide gel electrophoresis, HPLC analysis, and the like).

[0046] The term "expression cassette" refers to a nucleotide sequence that contains at least one coding sequence along with sequence elements that direct the initiation and termination of transcription. An expression cassette may include additional sequences, including but not limited to promoters, enhancers, and sequences involved in post-transcriptional or post-translational processes.

[0047] "Expression system" refers to a system for expressing a recombinant protein within a host cell. Generally, an expression system consists of a vector containing a genetic element encoding the protein to be expressed. This genetic element may comprise an expression cassette, which can include one or more elements for controlling expression, including: promoter sequences, transcription initiation sequences, enhancer sequences, selectable elements, and reporter genes. The vector used in an expression system can

consist of a plasmid or cosmid, a bacteriophage such as lambda phage or M13 phage, or an animal virus, such as retrovirus, lentivirus, adenovirus, herpes simplex virus (HSV), cytomegalovirus (CMV), adeno-associated virus (AAV), papillomavirus, and simian virus (SV40). As with the expression cassette, a vector utilized in an expression system may contain a variety of elements for controlling expression, including: promoter sequences, transcription initiation sequences, enhancer sequences, selectable elements, and reporter genes. In addition, this vector may include an origin of replication.

[0048] A "coding sequence" or "coding region" refers to a nucleic acid molecule having sequence information necessary to produce a gene product, when the sequence is expressed.

[0049] The term "operably linked" or "operably inserted" means that the regulatory sequences necessary for expression of the coding sequence are placed in a nucleic acid molecule in the appropriate positions relative to the coding sequence so as to enable expression of the coding sequence. This same definition is sometimes applied to the arrangement other transcription control elements (e.g. enhancers) in an expression vector.

[0050] Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, polyadenylation signals, terminators, and the like, that provide for the expression of a coding sequence in a host cell.

[0051] The terms "promoter", "promoter region", or "promoter sequence" refer generally to transcriptional regulatory regions of a gene, which may be found at the 5' or 3' side of the coding region, or within the coding region, or within introns. Typically, a promoter is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating

transcription of a downstream (3' direction) coding sequence. The typical 5' promoter sequence is bounded at its 3' terminus by the transcription initiation site and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence is a transcription initiation site (conveniently defined by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase.

[0052] A "vector" is a replicon, such as plasmid, phage, cosmid, or virus to which another nucleic acid segment may be operably inserted so as to bring about the replication or expression of the segment.

[0053] The term "selectable element" refers to a gene encoding a product that, when expressed, confers a selectable phenotype such as antibiotic resistance on a transformed cell.

[0054] The term "reporter gene" refers to a gene that encodes a product that is detectable by standard methods, either directly or indirectly.

[0055] The term "origin of replication" refers to a fixed location within a DNA sequence that serves as the startpoint for replication. DNA polymerase and other replicative factors bind at particular DNA sequences within the origin of replication. With regards to vectors, the origin of replication provides the ability for the vector to replicate autonomously, independent of the host chromosome.

[0056] The term "nucleic acid construct," "DNA construct" or "DNA expression construct" is sometimes used to refer to a coding sequence or sequences operably linked to

appropriate regulatory sequences and inserted into a vector for transforming a cell. This term may be used interchangeably with the term "transforming DNA". Such a nucleic acid construct may contain a coding sequence for a gene product of interest, along with a selectable marker gene and/or a reporter gene.

[0057] A "heterologous" region of a nucleic acid construct is an identifiable segment (or segments) of the nucleic acid molecule within a larger molecule that is not found in association with the larger molecule in nature. Thus, when the heterologous region encodes a mammalian gene, the gene will usually be flanked by DNA that does not flank the mammalian genomic DNA in the genome of the source organism. In another example, a heterologous region is a construct where the coding sequence itself is not found in nature (e.g., a cDNA where the genomic coding sequence contains introns, or synthetic sequences having codons different than the native gene). Allelic variations or naturally-occurring mutational events do not give rise to a heterologous region of DNA as defined herein.

[0058] The term "DNA construct", as defined above, is also used to refer to a heterologous region, particularly one constructed for use in transformation of a cell. A cell has been "transformed" or "transfected" or “transduced” by exogenous or heterologous DNA when such DNA has been introduced inside the cell. The transforming DNA may or may not be integrated (covalently linked) into the genome of the cell. In prokaryotes, yeast, and mammalian cells for example, the transforming DNA may be maintained on an episomal element such as a plasmid. With respect to eukaryotic cells, a stably transformed cell is one in which the transforming DNA has become integrated into a chromosome so that it is

inherited by daughter cells through chromosome replication. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA.

[0059] The term "in vivo delivery" involves the use of any gene delivery system, such as viral- and liposome-mediated transformation for the delivery and introduction of a HE1 therapeutic agent to the cells of a subject, while they remain in the subject.

[0060] As used herein, the term "transduction" is used to describe the delivery of DNA to eukaryotic cells using viral mediated delivery systems, such as adenoviral, AAV, retroviral, or plasmid delivery gene transfer methods. Preferably the viral mediated delivery system is targeted specifically to the cell, wherein delivery is sought. The production of targeted delivery systems is well known and practiced in the recombinant arts. A number of methods for delivering therapeutic formulations, including DNA expression constructs (as described further below), into eukaryotic cells are known to those skilled in the art. In light of the present disclosure, the skilled artisan will be able to deliver the therapeutic agents of the present invention to cells in many different but effective ways. For instance, the specificity of viral gene delivery may be selected to preferentially direct the HE1 gene to a particular target cell by using viruses that are able to infect particular cell types. Naturally, different viral host ranges will dictate the virus chosen for transfer.

[0061] "In vitro" gene delivery refers to a variety of methods for introducing exogenous DNA into a cell that has been removed from its host environment.

[0062] As used herein, the term "transfection" is used to describe the delivery and introduction of a therapeutic agent to a cell using non-viral mediated means, these methods include, e.g., calcium phosphate- or dextran sulfate-mediated transfection; electroporation; glass projectile targeting; and the like. These methods are known to those of skill in the art, with the exact compositions and execution being apparent in light of the present disclosure.

[0063] "Ex vivo gene delivery" refers to the procedure wherein appropriate cells are removed from an organism, transformed, transduced, or transfected in accordance with the teachings of the present invention, and replaced back into a host organism, for the purpose of therapeutic restoration and/or prevention.

[0064] "Delivery of a therapeutic agent" may be carried out through a variety of means, such as by using parenteral delivery methods such as intravascular and intramuscular injection, and the like. Such methods are known to those of skill in the art of drug delivery, and are further described herein in the sections regarding pharmaceutical preparations and treatment.

[0065] The term "contacted" when applied to a cell is used herein to describe the process by which an HE1 gene, protein or antisense sequence, and/or an accessory element (such as a an antibody or cytotoxic agent), is delivered to a target cell or is placed in direct proximity with the target cell. This delivery may be in vitro or in vivo and may involve the use of a recombinant vector system. Any method may be used to contact a cell with the HE1 associated protein or nucleotide sequence, so long as the method results in either increased or decreased levels of functional HE1 protein within the cell. This includes both

the direct delivery of an HE1 protein to the cell and the delivery of a gene or DNA segment that encodes HE1, or its antisense polynucleotide sequence, which gene or antisense sequence will direct or inhibit, respectfully, the expression and production of HE1 within the cell. Since protein delivery is subject to drawbacks, such as degradation and low cellular uptake, it is contemplated that the use of a recombinant vector that expresses a HE1 protein, or encodes for an HE1 polynucleotide or antisense sequence, will be of particular advantage for delivery.

[0066] The term "mammal" refers to such organisms as mice, rats, rabbits, goats, horse, sheep, cattle, cats, dogs and pigs. More preferably, "mammals" refers to monkeys and apes, and most preferably it refers to humans.

[0067] The phrase "therapeutically effective amount" describes an amount of the HE1 polynucleotide, antisense polynucleotide, peptide, protein, or portion thereof that is effective to bring about a desired effect when administered to a subject (e.g. an increase or decrease in cell cholesterol accumulation) within the subject.

Polynucleotides

[0068] The present invention provides novel compositions and methods for treating Niemann-Pick type C2 (NP-C2) disease, involving the administration of an HE1 gene, polynucleotide sequence, anti-sense sequence, polypeptide, protein or fragments thereof. The present inventor has determined that the HE1 protein is responsible for the defective egress of cholesterol from lysosomes that characterize subjects suffering from NP-C2. It has been discovered that the HE1 protein binds cholesterol in lysosomes and may be responsible for the transport of cholesterol from the lysosome to the various cellular

targets there by ensuring that free LDL-derived cholesterol can escape from the lysosomal compartment. This mechanism explains why NP-C2 cells, in which HE1 is inactive, exhibit large-scale cholesterol accumulation and has led to the development of the compositions and methods, herein described below, for the treatment of NP-C2 disease. The summary of the invention described above is non-limiting and other features and advantages of the invention will be apparent from the following detailed description.

[0069] The present invention concerns compositions and methods for treating various conditions related to the abnormal cellular accumulation of cholesterol, specifically that associated with the defective egress of cholestrol from lysosomes, namely NP-C2. The invention is based firstly on the inventor's discovery that HE1 mRNA and proteins were undetected in fibroblasts cells removed from subjects diagnosed with NP-C2; and that upon the administration of a therapeutic agent that includes an HE1 genetic element (e.g., an HE1 polynucleotide or peptide sequence), the diseased state could be alleviated.

[0070] The HE1 polynucleotides to be used by the present invention include isolated polynucleotides encoding HE1 polypeptides, proteins and fragments, and polynucleotides closely related thereto. The present invention identifies HE1 as the gene responsible for NP-C2 disease. The human HE1 was originally cloned by differential screening of a human epididymal cDNA library, and found to be highly expressed in all parts of the human epididymis (15, 16). The genomic structure of HE1 was determined by sequence alignments between HE1 cDNA (accession number Q15668) and genomic DNA sequence (accession number AC005479). The sequence of the human HE1 gene is set out in SEQ. ID. NO: 1. The HE1 gene encodes a 151 amino acid glycoprotein of 25-27 kDA. HE1 is

present in numerous cDNA and SAGE libraries (see UniGene cluster Hs.119529, www.ncbi.nlm.nih.gov/UniGene). A BLAST search indicates that the human HE1 protein shares extensive sequence homology with epididymal proteins from P. troglodyte and M. fascicularis.

[0071] One preferred embodiment of the present invention involves a pharmaceutical composition that includes polynucleotides of the human nucleotide sequences contained in SEQ ID NO: 1 encoding an HE1 polypeptide of SEQ ID NO:1. Compositions of the present invention further include an HE1 polynucleotide sequence comprising a nucleotide sequence that has at least 70% identity over its entire length to a nucleotide sequence encoding the HE1 polypeptide of SEQ ID NO: 1.

[0072] Hence, the nucleotide sequences encoding the HE1 polypeptide of SEQ ID NO:1 may be identical to the polypeptide encoding sequence contained in SEQ ID NO:2, or it may be a sequence, which as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO: 1. In this regard, polynucleotides with at least 70% identity are preferred, more preferably at least 80% even more preferably at least 90% identity, yet more preferably at least 95% identity are highly preferred and those with at least 98-99% are most highly preferred. Also included under HE1 polynucleotides are a nucleotide sequence that has sufficient identity to a nucleotide sequence contained in SEQ ID NO: 1 as to hybridize under conditions useable for amplification or for use as a probe or marker. The invention also provides polynucleotides that are complementary (e.g., antisense polynucleotide sequences) to such HE1 polynucleotides.

[0073] Also included in the present invention are polynucleotides encoding polypeptides which have at least 70% identity, preferably at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, even more preferably at least 97-99% identity, to the amino acid sequence of SEQ ID NO:1, over the entire length of the recited amino acid sequences.

[0100] HE1 polynucleotides (including antisense sequences) of the present invention may be prepared by two general methods: (1) they may be synthesized from appropriate nucleotide triphosphates, or (2) they may be isolated from biological sources. Both methods utilize protocols well known in the art. The availability of nucleotide sequence information, such as the cDNA having SEQ ID NO:2, enables preparation of compositions that include isolated HE1 nucleic acid molecules produced by oligonucleotide synthesis.

[0101] Compostions that include synthetic HE1 oligonucleotides may be prepared by the phosphoramadite method employed in the Applied Biosystems 38A DNA Synthesizer or similar devices. The resultant construct may be purified according to methods known in the art, such as high performance liquid chromatography (HPLC). A synthetic DNA molecule so constructed may then be cloned and amplified in an appropriate vector. HE1 genes also may be isolated from appropriate biological sources using methods known in the art.

[0102] For instance, HE1 may be isolated from genomic libraries of any mammal, specifically, human. A preferred means for isolating HE1 genes is PCR amplification using genomic or cDNA templates and HE1 specific primers. Genomic and cDNA libraries are commercially available, and can also be made by procedures well known in

the art. In positions of degeneracy where more than one nucleic acid residue could be used to encode the appropriate amino acid residue, all the appropriate nucleic acid residues may be incorporated to create a mixed oligonucleotide population, or a neutral base such as inosine may be used. The strategy of oligonucleotide design is well known in the art.

[0103] In accordance with the present invention, compositions of nucleic acids having the appropriate level of sequence homology (i.e., 70% identity or greater) with part or all the coding regions of SEQ ID NO:2 may be identified by using hybridization and washing conditions of appropriate stringency. For example, hybridizations may be performed, according to the method of Sambrook et al., using a hybridization solution comprising: 1.0% SDS, up to 50% formamide, 5x SSC (150mM NaCl, 15mM trisodium citrate), 0.05% sodium pyrophosphate (pH7.6), 5x Denhardt's solution, and 100 microgram/ml denatured, sheared salmon sperm DNA. Hybridization is carried out at 37-42°C for at least six hours. Following hybridization, filters are washed as follows: (1) 5 minutes at room temperature in 2X SSC and 1% SDS; (2) 15 minutes at room temperature in 2X SSC and 0.1% SDS; (3) 30 minutes to 1 hour at 37°C in 2X SSC and 0.1% SDS; (4) 2 hours at 45-55°C in 2X SSC and 0.1% SDS, changing the solution every 30 minutes.

[0104] One common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid molecules of a specified percent identity is set forth by (Sambrook et al., 1989, supra):

$$T_m = 81.5°C + 16.6 \mathrm{Log}\,[Na+] + 0.41\,(\%\ G\text{-}C) - 0.63\,(\%\ \mathrm{formamide}) - 600/\#\mathrm{bp\ in\ duplex}$$

As an illustration of the above formula, using [N+] = [0.368] and 50% formamide, with GC content of 42% and an average probe size of 200 bases, the T_m is 57°C. The T_m of a DNA duplex decreases by 1 - 1.5°C with every 1% decrease in homology. Thus, targets with greater than about 75% sequence identity would be observed using a hybridization temperature of 42°C.

[0105] The stringency of the hybridization and wash depend primarily on the salt concentration and temperature of the solutions. In general, to maximize the rate of annealing of the probe with its target, the hybridization is usually carried out at salt and temperature conditions that are 20 - 25°C below the calculated T_m of the of the hybrid. Wash conditions should be as stringent as possible for the degree of identity of the probe for the target. In general, wash conditions are selected to be approximately 12 - 20°C below the T_m of the hybrid. In regard to the nucleic acids of the current invention, a moderate stringency hybridization is defined as hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 μg/ml denatured salmon sperm DNA at 42°C, and wash in 2X SSC and 0.5% SDS at 55°C for 15 minutes. A high stringency hybridization is defined as hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 μg/ml denatured salmon sperm DNA at 42°C, and wash in 1X SSC and 0.5% SDS at 6-5°C for 15 minutes. Very high stringency hybridization is defined as hybridization in 6X SSC, SX Denhardt's solution, 0.5% SDS and 100 μg/ml denatured salmon sperm DNA at 42°C, and wash in 0. 1 X SSC and 0.5% SDS at 65°C for 15 minutes.

[0106] The nucleic acids to be used in the present invention may be maintained as DNA in any convenient cloning vector. In a preferred embodiment, clones are maintained in

plasmid cloning/expression vector, such as pBluescript (Stratagene, La Jolla, CA), that is propagated in a suitable *E. coli* host cell.

[0107] The above components can be utilized to generate a pharmaceutical composition consisting of an effective amount of a purified HE1 polynucleotide and an acceptable carrier. Preferably, this polynucleotide has at least 70% similarity with the sequence of SEQ. ID. NO: 1, and encodes a polypeptide sequence substantially similar to SEQ. ID. NO: 2. In addition, HE1 polynucleotides can comprise the genetic element that is incorporated into an expression vector. Such incorporation allows for the therapeutic expression of HE1 in a host cell, which can ameliorate certain disease states in which cholesterol transport and regulation is faulty. HE1 polynucleotides that can be utilized in this therapeutic expression system include HE1 DNA, cDNA, RNA, or antisense polynucleotides. Antisense HE1 polynucleotides can be utilized to downregulate HE1, which could be of therapeutic benefit in any disease in which there is excess lysosomal cholesterol transport.

Polypeptides

[0108] In one aspect, the present invention relates to pharmaceutical compositions that include human HE1 polypeptides (or HE1 proteins). The human HE1 polypeptides include the polypeptide of SEQ ID NO:1; as well as polypeptides comprising the amino acid sequence of SEQ ID NO:1; and polypeptides comprising the amino acid sequences which are substantially similar to the amino acid sequence of SEQ ID NO:1, that is they have at least 70% identity to that of SEQ ID NO:1, over its entire length. Preferably HE1 polypeptides exhibit at least one biological activity of HE1. The present invention further

provides for a pharmaceutical composition that includes a polypeptide that comprises an amino acid sequence which has at least 80% identity, more preferably at least 90% identity, yet more preferably at least 95% identity, most preferably at least 97-99% identity, to that of SEQ ID NO:2 over the entire length of SEQ ID NO:2.

[0109] The HE1 polypeptides, to be used in the compositions and methods of the present invention, may be in the form of the "mature" protein or may be a part of a larger protein such as a fusion protein or part of smaller fragments of the HE1 polypeptide that maintain functionality akin to the wild-type protein. A fragment is a polypeptide having an amino acid sequence that entirely is the same as part, but not all, of the amino acid sequence of the aforementioned HE1 polypeptides. Preferred fragments include, for example, truncation polypeptides having the amino acid sequence of HE1 polypeptides, except for deletion of a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Also preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions.

[0110] Other preferred fragments are biologically active fragments. Biologically active fragments are those that mediate HE1 activity, including those with a similar activity or an

improved activity, or with a decreased undesirable activity. Variants of the defined sequence and fragments also form part of the present invention. Preferred variants are those that vary from the referents by conservative amino acid substitutions.

[0111] The HE1 proteins and polypeptides to be used in the compositions and methods of the present invention can be isolated, purified and prepared in any suitable manner, such as those well known in the art. For instance, the production of the proteins using in vitro expression is a method well known in the art. For example, a cDNA or gene may be cloned into an appropriate in vitro transcription vector, for in vitro transcription, followed by cell-free translation in a suitable cell-free translation system. In vitro transcription and translation systems are commercially available, e.g., from Promega Biotech, Madison, Wisconsin, or BRL, Rockville, Maryland.

[0112] According to a preferred embodiment, larger quantities of HE1 encoded polypeptide, for use with the present invention, may be produced by expression in a suitable prokaryotic or eukaryotic system. For example, part or all of a HE1 DNA molecule, such as the coding portion of SEQ ID NO:2 may be inserted into a plasmid vector adapted for expression in a bacterial cell (such as *E. coli*) or a yeast cell (such as *Saccharomyces cerevisiae),* or into a baculovirus vector for expression in an insect cell. Such vectors comprise the regulatory elements necessary for expression of the HE1 DNA in the host cell, positioned in such a manner as to permit expression of the DNA into the host cell. Such regulatory elements required for expression include promoter sequences, transcription initiation sequences and, optionally, enhancer sequences. Plasmids specifically designed to express and secrete foreign proteins are available from

commercial sources. For example, if expression is desired in *E. coli*, commonly used plasmids include pTrcPPA (Pharmacia); pPROK-C and pKK233-2 (Clontech); and pNH8a, pNH16a, pcDNAII and pAX (Stratagene), among others.

[0113] The HE1 proteins produced by in vitro transcription and translation or by gene expression in a recombinant procaryotic or eukaryotic system may be purified according to methods known in the art. Recombinant proteins can be substantially purified by affinity separation, such as by immunological interaction with antibodies that bind specifically to the recombinant protein or fusion proteins such as His tags. Such methods are commonly used by skilled practitioners. The HE1 proteins thus prepared, may then be analyzed according to standard procedures. For example, the protein may be subjected to amino acid composition, amino acid sequence, or protein concentration analysis according to known methods.

[0114] Using appropriate amino acid sequence information, synthetic HE1 proteins of the present invention may be prepared by various synthetic methods of peptide synthesis via condensation of one or more amino acid residues, in accordance with conventional peptide synthesis methods. Preferably, peptides are synthesized according to standard solid-phase methodologies, such as may be performed on an Applied Biosystems Model 430A peptide synthesizer (Applied Biosystems, Foster City, CA), according to manufacturer's instructions. Other methods of synthesizing peptides or peptidomimetics, either by solid phase methodologies or in liquid phase, are well known to those skilled in the art.

[0115] In particular aspects, the invention relates to compositions and methods for using such polypeptides and polynucleotides for treating diseases associated with increased

cellular levels of cholesterol, for instance NP-C2 disease, by administering a HE1 gene or protein, in a pharmaceutically acceptable and appropriate delivery vehicle, to increase HE1 mediated cholesterol egress from lysosomes and the down-regulation of endogenous cellular production of cholesterol. Further, the compositions and methods of the present invention may be used for treating a disease associated with increased lysosomal cholesterol transport by administering an HE1 antisense polynucleotide sequence in a pharmaceutically acceptable and appropriate delivery vehicle.

[0116] The pharmaceutical compositions of the present invention may be formulated and used as tablets, capsules, or elixirs for oral administration; suppositories for rectal or vaginal administration; sterile solutions and suspensions for parenteral administration; creams, lotions, or gels for topical administration; aerosols or insufflations for intratracheobronchial administration; and the like. Preparations of such formulations are well known to those skilled in the pharmaceutical arts. The dosage and method of administration can be tailored to achieve optimal efficacy and will depend on factors that those skilled in the medical arts will recognize.

[0117] When administration is to be parenteral, such as intravenous on a daily basis, injectable pharmaceuticals may be prepared in conventional forms, either as liquid solutions or suspensions; solid forms suitable for solution or suspension in liquid prior to injection; or as emulsions. Suitable excipients are, for example, water, saline, dextrose, mannitol, lactose, lecithin, albumin, sodium glutamate, cysteine hydrochloride, or the like. In addition, if desired, the injectable pharmaceutical compositions may contain minor amounts of nontoxic auxiliary substances, such as wetting agents, pH buffering agents,

and the like. If desired, absorption enhancing preparations (e.g. liposomes) may be utilized.

[0118] For administration, a therapeutic agent consisting of an HE1 polypeptide or a system capable of expressing this polypeptide, will generally be mixed prior to administration with a non-toxic, pharmaceutically acceptable carrier substance. Usually, this will be an aqueous solution, such as normal saline or phosphate-buffered saline (PBS), Ringer's solution, lactate-Ringer's solution, or any isotonic physiologically acceptable solution for administration by the chosen means. Preferably, the solution is sterile and pyrogen-free, and is manufactured and packaged under current Good Manufacturing Processes (GMP's), as approved by the FDA. The clinician of ordinary skill is familiar with the appropriate ranges for pH, tonicity, and additives or preservatives when formulating pharmaceutical compositions for administration by intravascular injection, intrathecal injection, direct injection into aberrant cells, or by other routes.

[0119] The effective amount of the therapeutic composition to be given to a particular patient will depend on a variety of factors, several of which will be different from patient to patient. A competent clinician will be able to determine an effective amount of therapeutic composition to administer to a patient to effect appropriate cholesterol regulation within a subject's cells. Dosage of the therapeutic composition will depend on the treatment of the particular disease (e.g., such as NP-C2), route of administration, the nature of the therapeutic delivery vehicle, etc. Utilizing LD_{50} animal data, and other information available for such administrations, a clinician can determine the maximum safe dose for a subject, depending on the route of administration. For instance, an

intravenously administered dose may be more than intrathecally administered dose, given the greater body of fluid into which the therapeutic composition is being administered. Similarly, compositions, which are rapidly cleared from the body, may be administered at higher doses, or in repeated doses, in order to maintain a therapeutic concentration. Utilizing ordinary skill, the competent clinician will be able to optimized the dosage of a particular therapeutic composition in the course of routine clinical trials.

[0120] The above components can be utilized to create a pharmaceutical composition consisting of an HE1 polypeptide and an acceptable carrier. Preferably, this HE1 polypeptide is substantially similar to SEQ. ID. NO:1, or to a fragment thereof. The HE1 polypeptide can be administered directly in the form of purified recombinant protein, to a subject suffering from NP-C2 disease, or it can be expressed within a host cell by presenting the cell with an expression vector containing the HE1 polynucleotide sequence. Administration of the HE1 polypeptide can restore normal cholesterol transport and regulation within a cell by restoring cholesterol egress from the lysosome.

Vectors, Host Cells, and Expression

[0121] The invention also relates to pharmaceutical compositions that include expression systems that contain expression cassettes on vectors that comprise a HE1 polynucleotide, or polynucleotides for use in recombinant techniques involving both in vitro and in vivo, as well as ex vivo gene therapy procedures. These expression systems can be utilized to express an HE1 polypeptide or polypeptide fragment, HE1 RNA sequence, or HE1 antisense RNA sequence within a host cell. Expression systems can be utilized to express

HE1 polypeptide sequences in either prokaryotic and eukaryotic cells, for both investigative and therapeutic purposes.

[0122] Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides and polypeptides into host cells can then be effected by methods described in many standard laboratory manuals, such as Davis et al., *Basic Methods In Molecular Biology (1986)* and *Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)* such as calcium phosphate transfection, DEAE-dextran mediated transfection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

[0123] Representative examples of appropriate hosts for in vitro procedures include bacterial cells, such as streptococci, staphylococci, *E. coli, Streptomyces* and Bacillus subtilis cells; fungal cells, such as yeast cells and Aspergiffits cells, insect cells such as *Drosophila* S2 and *Spodoptera Sf9* cells; animal cells such as Fibroblasts, CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells, and plant cells. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

[0124] More particularly, the present invention also includes pharmaceutical compositions comprising recombinant constructs that include a HE1 DNA, cDNA or RNA sequence.

Such a construct comprises a vector, such as a plasmid or viral vector, into which the clone has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the genetic sequence, transcription initiation sequences and enhances sequences. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example; Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBS, pD10, phagescript, psiX 174, pbluescript SK, pbsks, pNH8A, pNH 16a, pNHI8A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). As further examples, cDNA of human HE1 may be inserted in the pEF/myc/cyto vector (from Invitrogen) and/or the pCMV-Tag3b vector (from Stratagene), which can then be used with anti-Myc Ab, to transform Stem, HeLa, Fibroblast (or other) cells with HE1 DNA. Or the HE1 protein may be isolated, purified and directly injected in to the appropriate tissue, infused to blood cells, or delivered in a lyophilized carrier as described above.

[0125] However, any other plasmid or vector may be used as long as they are replicable and viable in the host's cells. In addition, a complete mammalian transcription unit and a selectable marker can be inserted into a prokaryotic plasmid for use in in vivo or ex-vivo procedures. The resulting vector is then amplified in bacteria before being transfected into cultured mammalian cells or delivered directly to the subject with an acceptable biological carrier as described below. Examples of vectors of this type include pTK2, pHyg and

pRSVneo. Hence, these plasmids, constructs, cassettes and vectors may be used in both in vivo and ex vivo procedures. As stated, ex vivo procedures involve the removal of a host cell, such as a fibroblast cell, from the subject, recombinant manipulation of the cell (i.e., transformation, transduction or transfection with a suitable HE1 expression system vector), and the re-delivery of the cell back into a host's environment.

[0126] Further, according to one particular embodiment of the present invention, recombinant HE1 DNA, cDNA or RNA may be directly injected to fibroblasts for the production of HE1 endogenously. While, the polynucleotide sequence coding for the antisense sequence encoding the protein RNA, may be directly injected to fibroblasts for the endogenous inhibition of HE1. HE1 DNA, cDNA, RNA or polynucleotide sequences coding for the antisense sequence encoding the protein may also be delivered using other appropriate means, including vectors, as described, and well known in the recombinant arts.

[0127] A wide variety of recombinant plasmids may be engineered to express the HE1 protein and used to deliver HE1 to a cell. These include the use of naked DNA and HE1 plasmids to directly transfer genetic material into a cell (Wolfe et al., 1990); formulations of HE1 encoding trapped liposomes (Ledley et. al., 1987) or in proteoliposomes that contain other viral envelope receptor proteins (Nicolau et al., 1983); and HE1-encoding DNA, or antisense sequence, coupled to a polysineglycoprotein carrier complex. The HE1 encoding sequence may be incorporated as part of an expression cassette, a nucleotide sequence that contains at least one coding sequence along with sequence elements that direct the initiation and termination of transcription. An expression cassette may include

additional sequences, including but not limited to promoters, enhancers, and sequences involved in post-transcriptional or post-translational processes. Hence methods for the delivery of nucleotide sequences to cells are well known in the recombinant arts. Such methods for in vitro delivery, further include, but are not limited to: microinjection, calcium phosphatase, lyposomes, and electroporation.

[0128] Genetic material, such as HE1 nucleotides of the present invention, may be delivered to cells, in vivo or ex vivo, using various different plasmid based delivery platforms, including but not limited to recombinant ADV (such as that described in U.S. Pat. No. 6,069,134 incorporated by reference herein), AAV (such as those described by U.S. Pat. No. 5,139,941 incorporated by reference herein), MMLV, Herpes Simplex Virus (U.S. Pat. No. 5,288,641, incorporated by reference herein), cytomegalovirus, lentiviral, and overall, retroviral gene delivery systems, well known and practiced with in the art.

[0129] Techniques for preparing replication defective, infective viruses are well known in the art, as exemplified by Ghosh-Choudhury & Graham (9187); McGory et al. (1988); and Gluzman et al. (1982), each incorporated by reference herein. These systems typically include a plasmid vector including a promoter sequence (such as CMV early promoter) operably linked to the nucleotide coding the gene of interest (inserted into an appropriate gene insertion site; i.e., an IRES site), as well as transcription initiation sequences, enhancer sequences, a terminating signal (such as a Poly-A tail i.e., BGH), and the appropriate mutations so as to make the delivery vehicle replication defective (e.g., Psi sequence deletions) and safe for therapeutic uses. The construction of the appropriate

elements in a vector system containing the nucleotides of the present invention is well within the skills of one versed in the recombinant arts.

[0130] A great variety of vector and/or expression systems can be used. Such systems include, among others, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia, viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any delivery system (such as a vector) suitable to maintain, propagate or express polynucleotides to produce a polypeptide in a host may be used. The appropriate HE1 nucleotide sequence may be inserted into delivery system by any of a variety of well-known and routine techniques, such as, for example, those set forth in *Sambrook et al., Molecular Cloning, A Laboratory Manual (supra).*

[0131] Promoter regions can be selected from any desired gene using CAT (chloramphenicol acetyl transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, PL and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-1. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

[0132] Administration of the HE1 polypeptide, or a vector capable of expressing the HE1 polypeptide, can be utilized to treat diseases involving faulty cholesterol transport and regulation. A vector containing an HE1 genetic element can be used in conjunction with a biologically acceptable carrier to express HE1 in a host cell. Preferably, such the HE1 genetic element in such an expression system comprises a polynucleotide sequence with at least 70% similar to SEQ. ID. NO: 1. However, another HE1 genetic element, such as HE1 RNA, cDNA, or antisense polynucleotide can be used instead. Preferably, the HE1 genetic element is contained within an expression cassette. Administration of the HE1 expression system to a cell can serve to restore both lysosomal cholesterol transport and endogenous cholesterol regulation. Expression of HE1 within a host cell can thus potentially ameliorate not only NP-C2 disease, but any disease in which cholesterol regulation is faulty.

[0133] If the HE1 polypeptide is to be expressed for use in screening assays, generally, the protein is produced intracellularly requiring the cells first to be lysed before the polypeptide is recovered. The HE1 polypeptides can be recovered and substantially purified from recombinant cell cultures by well known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

[0134] HE1 screening assays can be utilized to screen for both NP-C2 and NP-C1 diseases. Patients with NP-C2 disease have either a homozygous or heterozygous mutation in both copies of the gene encoding for HE1. Detection of a mutation in the HE1 gene sequence of SEQ. ID. NO: 1 can be utilized to diagnose NP-C2 disease. This test can be performed prenatally by analyzing the HE1 sequence of a fetus. HE1 sequence analysis can also be utilized to diagnose the ability of a subject to genetically transmit NP-C2 disease. A subject with a mutation in only one copy of the HE1 gene will exhibit no signs of the disease, but would possess a higher risk of genetically transmitting the disease to their offspring. Mutations in the HE1 gene can consist of deletions, substitutions, splice mutations, or insertions. Mutations can be detected by PCR amplification of a subject's DNA, followed by sequence analysis of the HE1 gene. Alternatively, mutations in the HE1 gene can be detected by hybridizing the HE1 gene sequence of a subject with a labeled polynucleotide probe complementary to the HE1 gene sequence of SEQ. ID. NO: 1. Generally, this polynucleotide probe will consist of a DNA molecule bound to a label selected from radioactive isotopes (e.g. ^{32}P, ^{125}I, ^{35}S), biotin, an enzyme reporter group (e.g. horseradish peroxidase, alkaline phosphatase), a chemiluminescent label, a fluorescent label, or an antibody. To increase the sensitivity of the probe, the target HE1 gene sequence may be amplified by PCR. Target sequence and probe may be hybridized on a solid support such as a filter membrane, in solution, in situ, or via Southern blot. A battery of probes may be utilized to ensure hybridization between target and probe for any potential mutation.

[0135] In addition to NP-C2, HE1 screening can serve as a method of diagnosing NP-C1 disease. Patients with NP-C1 exhibit raised levels of HE1 expression. Thus, a comparison of HE1 protein or mRNA levels in a cell as compared to a healthy control cell can be used to diagnose the presence of NP-C1 disease. Increased HE1 expression can be detected by measuring HE1 mRNA levels using techniques such as Northern blot, RT-PCR, or ribonuclease protection assay, or by measuring HE1 protein levels using techniques such as Western blotting.

[0136] The above-described constructs, cassettes, plasmids, and vectors are useful in gene therapy procedures. Successful gene therapy generally requires the integration of a gene able to correct the genetic disorder into the host genome, where it would co-exist and replicate with the host DNA and be expressed at a level to compensate for the defective gene. Ideally, the disease would be cured by one or a few treatments, with no serious side effects. There are several approaches to gene therapy proposed.

[0137] As described above, basic transfection methods exist in which DNA containing the gene of interest is introduced into cells non-biologically, for example, by permeabilizing the cell membrane physically or chemically. Liposomes or protein conjugates formed with certain lipids and amphophilic peptides can be used for transfection. (Stewart et al., 1992; Torchilin et al., 1992; Zhu et al., 1993, incorporated herein by reference.) This approach is particularly effective in ex vivo procedures involving fibroblasts, which can be temporarily removed from the body and can tolerate this mode of treatment.

[0138] A second, transduction approach, capitalizes on the natural ability of viruses to enter cells, bringing their own genetic material with them. For example, Retroviruses

have promise as gene delivery vectors due to their ability to integrate their genes into the host genome, transferring a large amount of foreign genetic material, infecting a broad spectrum of species and cell types and of being packaged in special cell-lines (Miller, 1992, incorporated herein by reference).

[0139] A third method uses other viruses, such as adenovirus, herpes simplex virues (HSV), cytomegalovirus (CMV), and adeno-associated virus (AAV), which are engineered to serve as vectors for gene transfer. Although some viruses that can accept foreign genetic material are limited in the number of nucleotides they can accommodate and in the range of cells they infect, these viruses have been demonstrated to successfully effect gene expression. For example, adenovirus gene transfer systems may be used. Such a system is based upon recombinant, engineered adenovirus which is rendered replication-incompetent by deletion of a portion of its genome, such as E1, and yet still retains its competency for infection. Relatively large foreign proteins can be expressed when additional deletions are made in the adenovirus genome. For example, adenoviruses deleted in both E1 and E3 regions are capable of carrying up to 10 Kb of foreign DNA and can be grown to high titers in 293 cells (Stratford-Perricaudet and Perricaudet, 1991a). Surprisingly persistent expression of transgenes following adenoviral infection has also been reported.

[0140] Further, the recombinant HE1 DNA, cDNA or RNA may be delivered to cells, for the production of HE1 endogenously, and the polynucleotide sequence coding for the antisense sequence encoding the protein RNA, may be delivered to cells, for the endogenous inhibition of HE1, by use of biologically compatible carriers or excipients.

This may be useful in inducing or inhibiting intracellular cholesterol levels and cholesterol transport. Pharmaceutically acceptable carriers for therapeutic use are well known in the pharmaceutical art, and are described, for example, in Remington's Pharmaceutical Sciences (A. P. Gennaro, ed.; Mack, 1985). For example, sterile saline or phosphate-buffered saline at physiological pH may be used. Preservatives, stabilizers, dyes, and even flavoring agents may be provided in the pharmaceutical composition. For example, sodium benzoate, sorbic acid, and esters of p-hydroxybenzoic acid may be added as preservatives. Antioxidants and suspending agents may also be used. As stated above, the dosage and method of administration can be tailored to achieve optimal efficacy and will depend on factors that those skilled in the medical arts will recognize. When administration is to be parenteral, such as intravenous on a daily basis, injectable pharmaceuticals may be prepared in conventional forms, either as liquid solutions or suspensions; solid forms suitable for solution or suspension in liquid prior to injection; or as emulsions. Suitable excipients are, for example, water, saline, dextrose, mannitol, lactose, lecithin, albumin, sodium glutamate, cysteine hydrochloride, or the like. In addition, if desired, the injectable pharmaceutical compositions may contain minor amounts of nontoxic auxiliary substances, such as wetting agents, pH buffering agents, and the like. If desired, absorption enhancing preparations (e.g. liposomes) may be utilized.

[0141] As described above and exemplified below, a preferred embodiment of the present invention involves methods for the treatment of diseases associated with an unhealthy increase or decrease of intracellular levels of cholesterol in a subject. These methods

involve administering to the subject a pharmaceutical composition that includes an effective amount of a HE1 protein or a nucleotide sequence coding for the HE1 protein or a nucleotide sequence that codes for the anti-sense sequence of the nucleotide sequence coding for the HE1 protein. These may be delivered by suitable means, as described above, including the use of vectors and or acceptable biological carriers.

[0142] The foregoing is intended to be illustrative of the embodiments of the present invention, and are not intended to limit the invention in any way. Although the invention has been described with respect to specific modifications, the details thereof are not to be construed as limitations, for it will be apparent that various equivalents, changes and modifications may be resorted to without departing from the spirit and scope thereof and it is understood that such equivalent embodiments are to be included herein. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Example 1: Analysis of human brain mannose 6-phosphorylated glycoproteins

[0143] Identification of HE1 as the gene responsible for NP-C2 disease arose from an ongoing study directed at characterizing the lysosome proteome. The approach of this study was based on the fact that many soluble lysosomal proteins acquire a post-translational modification that distinguishes them from most other types of proteins, the mannose 6-phosphate (Man6-P) marker. This modification is recognized by Man6-P receptors (MPRs), which divert newly synthesized lysosomal enzymes from the secretory

pathway to the endolysosomal system (33). Purified MPR derivatives typically bind phosphorylated lysosomal proteins with subnanomolar affinity and can be used to detect and purify Man6-P glycoproteins (32, 34).

[0144] MPR affinity-purified proteins from the human brain were fractionated by two-dimensional electrophoresis with precast gels (Invitrogen, Carlsbad, CA). The first dimension consisted of isoelectric focusing on pH 3 to 10 gels without prior sample reduction, while the second dimension consisted of sample reduction by SDS-PAGE on 10 to 20% polyacrylamide gels. The resultant two-dimensional gel map of MPR affinity-purified proteins contained a group of proteins sharing the same NH_2-terminal sequence, as determined by Edman degradation, with these sequences likely representing differentially glycosylated isoforms of the same protein. The longest of these sequences consisted of the amino acid sequence EPVQFKDXGSVDGVIK, which matches perfectly the processed NH_2-terminus of HE1. The group of proteins sharing this NH_2-terminal sequence is circled in the Coomassie blue-stained polyvinylidine difluoride membrane in Figure 1A.

[0145] Based on this amino acid sequence similarity, Western blot experiments were run to confirm that the MPR affinity-purified proteins did indeed represent differentially glycosylated versions of HE1. COOH-terminally hexahistidine-tagged HE1 was expressed in *Escherichia coli*. The expressed protein formed inclusion bodies, which were solubilized in guanidine-HCl. The protein was then purified by chromatography on immobilized cobalt, renatured by dilution, and further purified by anion exchange chromatography. Rabbit antiserum was raised against purified HE1, and this antiserum

was used to generate polyclonal antibodies for Western blotting. A nitrocellulose membrane containing 5μg of the MPR affinity-purified proteins, pictured in Figure 1B, was probed with these antibodies. These Western blot results confirmed that the MPR affinity-purified proteins containing the HE1 NH_2-terminus sequence were indeed HE1.

[0146] Figure 1C again illustrates a nitrocellulose membrane containing 5μg of the MPR affinity-purified proteins, this time probed with radiolabeled MPR. Probing with this 2 nM ^{125}I-labeled soluble cation-independent MPR confirmed that HE1 contained the Man6-P modification. Inclusion of 10 mM Man6-P during probing abrogated the signal.

[0147] Northern blot analysis was performed to measure the level of HE1 expression in various human tissues. 2 μg of polyadenylated RNA from various human tissues (Origene Technologies, Rockville, MD) was immobilized on a membrane and probed with a ^{32}P-labeled DNA fragment corresponding to the entire coding region of HE1. This Northern blot, pictured in Figure 1D, revealed a single transcript of 0.9 kb in all tissues examined. The highest levels of HE1 expression were observed in testis, kidney, and liver, while the lowest levels of expression were observed in lung and muscle.

Example II: Subcellular distribution of HE1 homolog in rat liver

[0148] Rat liver homogenate was subjected to differential centrifugation as described in (39). Differential centrifugation resulted in the homogenate differentiation into five fractions: N, the nuclear fraction; M, the heavy mitochondrial fraction; L, the light mitochondrial fraction; P, the microsomal fraction; and S, the soluble fraction. Each fraction was assayed for the presence of both HE1 and lysosomal protein activity. Figure 2A illustrates the results of these assays. The upper panel represents Western blot analysis

of each fraction with rabbit antibodies to HE1. After differential centrifugation, HE1 was found mainly in the heavy mitochondrial (M) fraction and the light mitochondrial (L) fraction. The lower panel of Figure 2A illustrates lysosomal protein activity in each fraction. The filled bars represent ß-galactosidase activity, the open bars represent tripeptidyl peptidase activity, and the hatched bars represent ß-hexosaminidase activity. The highest levels of lysosomal protein activity were observed in the heavy mitochondrial (M) fraction and the light mitochondrial (L) fraction. The presence of HE1 only in the fractions with the highest lysosomal protein activity showed that HE1 itself is a lysosomal protein.

[0149] The heavy mitochondrial (M) fraction and the light mitochondrial (L) fraction from the differential centrifugation experiment were pooled for analysis by isopycnic centrifugation as described in (41). The density gradient for isopycnic centrifugation was made of either sucrose or Nycodenz. Figure 2B shows the results of these isopycnic centrifugation experiments using fractions pooled from control rats. The dashed line represents the sucrose gradient. The upper panel represents Western blot analysis of each sucrose gradient fraction using rabbit antibodies to HE1. The lower panel represents lysosomal protein activity. As with the differential centrifugation results, there was a clear codistribution of HE1 and the three lysosomal enzymes, showing that they reside in the same compartment.

[0150] Differential centrifugation was then performed on liver homogenate taken from rats treated with Triton WR1339, a nonionic detergent that accumulates in lysosomes and induces a striking and specific shift of these organelles in sucrose density gradients (35).

The heavy mitochondrial (M) and light mitochondrial (L) fractions were again subjected to isopycnic centrifugation as described in (41), with a gradient made of either sucrose or Nycodenz. Figure 2C shows the results of these isopycnic centrifugation experiments using fractions pooled from treated rats. The dashed line represents the sucrose gradient. The upper panel represents Western blot analysis of each sucrose gradient fraction using rabbit antibodies to HE1, while the lower panel represents lysosomal protein activity. Codistribution of HE1 and lysosomal protein activity was once again observed, with both shifting to earlier fractions in the sucrose gradient. The results of these isopycnic centrifugation experiments verify that most, if not all, HE1 is located in lysosomes.

Example III: Measurement of HE1 levels in control and mutant fibroblasts

[0151] Western blots were performed to determine the level of HE1 in control and mutant fibroblasts. Fibroblasts originated from either the Coriell collection (designated by GM numbering) or from Peter Pentchev (designated by NPC numbering). Control fibroblasts were from cells lines GM06556, GM05757B, and GM03625F. Mutant fibroblasts originated from subjects having either sea blue histiocyte disease, NP-C1 disease, or NP-C2 disease. Sea blue histiocyte fibroblasts were from cell lines GM01912 and GM00843. NP-C1 disease fibroblasts were from cell lines GM11095, GM03123A, GM00110B, and NPC1 90.48. NP-C2 disease fibroblasts were from cell lines NPC2 93.10 and NPC2 99.04. Fibroblasts were cultured in RPMI 1640 medium supplemented with 15% fetal bovine serum. The cells were lysed in 0.1% Triton X-100 and 150 mM NaCl, and extracts were centrifuged 20 minutes at 13,000g. Soluble protein (7 μg) was separated by SDS-polyacrylamide gel electrophoresis. The blot was probed with rabbit antibodies to HE1

and chemiluminescence. As a control, the blot was reprobed with rabbit antibodies to cathepsin D (Calbiochem, LaJolla, CA) and chemiluminescence. The results of these blots are pictured in Figure 3A. The top panel represents Western blot analysis of HE1, while the lower panel represents Western blot analysis of the cathepsin D control. Lanes 1, 2, and 3 represent control fibroblast samples, lanes 4 and 5 represent sea blue histiocyte disease samples, lanes 6, 7, 8, and 9 represent NP-C1 disease samples, and lanes 10 and 11 represent NP-C2 disease samples. As seen in the top panel, HE1 expression was undetectable in fibroblasts from the NP-C2 patients. In contrast, HE1 expression was detectable in control fibroblasts as well as fibroblasts taken from patients with sea blue histiocyte disease and NP-C1 disease.

Example IV: Sequence analysis of HE1 gene in NP-C2 patients

[0152] The complete sequence of the human HE1 gene is available under GenBank accession number AC005479. A schematic of the HE1 gene and protein is presented in Figure 3B. In this schematic, nt1 represents the first nucleotide of the initiation codon. The arrowheads represent potential N-linked glycosylation sites, with the filled arrowhead denoting a site that is conserved among mammalian HE1 orthologs. The predicted disulfide pairing of the cysteines portrayed in the schematic is assigned by homology to equivalent cysteines in an apparently related dust mite protein (42). The genomic structure of HE1 was determined by sequence alignments between the HE1 cDNA (Q15668) and genomic DNA sequence (AC005479). The chromosomal localization of HE1 (14q24.3) was determined by identifying mapped clones in the Sequence Tagged Sites (STS) database that aligned to AC005479 (e.g. STS clones G38283, G38146, and

G38077) and was confirmed by radiation hybrid panel mapping to chromosome 14 with the Coriell monochromosomal somatic hybrid panel.

[0153] The human HE1 gene sequence was used to design polymerase chain reaction primers for amplifying the entire HE1 coding region. Sequence analysis of the amplified HE1 gene sequence revealed the presence of mutations in two unrelated NP-C2 patients. The location of these two mutations is designated by the dark arrows above the protein schematic in Figure 3B. One patient (NPC2 99.04) was homozygous for a transversion of G to T in exon 1 that results in conversion of amino acid E20 to a termination codon. Since E20 corresponds to the NH_2-terminus of the mature protein, this represents a null mutation. The other patient (NPC2 93.10) was compound heterozygous for the Glu20Stop mutation and a single nucleotide deletion in exon 2 that shifts the reading frame and generates a stop codon four codons downstream. This severe truncation also affects the predicted disulfide pairing of the protein and is likely to represent a null allele. No mutations were detected in sequence analysis of DNA from eight individuals who represented either unaffected controls, patients with NP-C1, or patients with other diseases. These findings illustrate that mutations in HE1 are specifically associated with NP-C2.

V. Correction of cholesterol accumulation in NP-C2 fibroblasts

[0154] The finding that NP-C2 is caused by a deficiency in the lysosomal HE1 protein led to experiments in which HE1 protein was added to NP-C2 fibroblasts in trans to reverse the cholesterol accumulation phenotype.

[0155] A fragment corresponding to human HE1 cDNA flanked by *Xho*I sites was generated using standard PCR-based methods and subcloned into *Xho*I-digested

expression vector pMSXND1, yielding a construct that contains an HE1 expression cassette, a neomycin-resistance cassette for G418 selection, and a dihydrofolate reductase expression cassette for MTX-based selection. After restriction mapping, correctly oriented constructs were sequenced to confirm absence of unwanted changes in the coding region. Plasmid DNA was linearized with *Pvu*I before transfection.

[0156] Chinese hamster ovary (CHO) cells were maintained in Dulbecco's modified Eagle's (DME)/F12 medium (Sigma) supplemented with 10% (vol/vol) fetal-bovine serum (FBS) and were stably transfected using LIPOFECTAMINE™ (Gibco). Cells were selected with G418 and resistant clones isolated using cloning cylinders. A resistant clone that secreted mannose 6-phosphorylated HE1 was subjected to multiple cycles of selection with increasing concentrations of MTX. Cells resistant to 400 μM MTX were subcloned and the clone that secreted the highest level of HE1 was chosen for subsequent analysis. Figure 4D represents a Western blot comparing HE1 expression in the HE1-transfected CHO cells (lane 2) to HE1 expression in normal CHO cells (lane 1). HE1 secreted from transfected CHO cells is secreted as a functional protein. Presumably, the overexpressed HE1 overwhelms the sorting machinery of the cell, resulting in secretion rather than lysosomal targeting. The secreted HE1 contains the mannose 6-phosphate modification, meaning it can be delivered by receptor mediated endocytosis and subsequently targeted to the lysosome.

[0157] Control NP-C2 fibroblasts were cultured for 4 days in RPMI 1640 medium (Gibco) supplemented with 15% FBS and maintained at 37 °C in a humidified atmosphere containing 5% CO_2. Experimental cells were cultured in the same manner, but were

treated with a supplement consisting of either 0.3% conditioned medium from the HE1 transfected CHO cells or 0.3% conditioned medium from untransfected CHO cells. Accumulation of cholesterol in the NP-C2 fibroblasts was then demonstrated by punctate fluorescence after probing with filipin, a cholesterol antibiotic. Figure 4A illustrates filipin staining results for control NP-C2 fibroblasts, while Figure 4B and 4C illustrate filipin staining results for NP-C2 fibroblasts treated 0.3% medium from HE1 producing CHO cells or normal CHO cells, respectively. As these results show, cultivation of NP-C2 cells with a small amount (0.3% vol/vol) of medium from the CHO cells transfected with HE1 diminished cholesterol accumulation compared with controls (naïve medium or equivalent amounts of conditioned medium from untransfected CHO cells). Cultivation of NP-C2 cells with large amounts (>10%) of conditioned media from untransfected CHO cells partially reversed cholesterol accumulation, presumably reflecting the presence of low amounts of endogenous HE1 homolog secreted by the CHO cells. Addition of Man6-P to the HE1-conditioned media prevented the reduction in cholesterol accumulation, indicating that uptake occurred through MPR-mediated endocytosis. In contrast to the results obtained with NP-C2 cells, comparable experiments showed that the HE1-conditioned medium had no effect on reducing cholesterol accumulation in NP-C1 fibroblasts, thus further demonstrating the specificity of the defect in NP-C2.

[0158] For quantitation, the NP-C2 fibroblasts were grown to confluence, and fields containing approximately 200 cells were selected under bright-field illumination to eliminate operator bias. Fluorescence measurements were then collected with charge-coupled device camera. For each condition, the average pixel intensity of five fields was

used for the analysis. Data were corrected for background staining obtained with unaffected control fibroblasts (477 ± 15, mean ± standard error). These results are pictured in Figure 4E, with A representing control NP-C2 fibroblasts, B representing NP-C2 fibroblasts treated with medium from HE1 producing CHO cells, and C representing NP-C2 fibroblasts treated with medium from normal CHO cells. The asterisk indicates that the difference in staining intensity of the HE1-treated fibroblasts compared with the two control groups is statistically significant ($P < 0.05$).

[0159] The finding that NP-C2 disease is due to a deficiency in a soluble lysosomal protein is consistent with earlier observations that cocultivation of mononuclear NP-C1 and NP-C2 fibroblasts partially reversed cholesterol accumulation in a subset of the cells (11). It may also explain why fibroblasts from patients with I-cell disease, which lack the enzyme that normally generates the Man6-P lysosomal targeting signal and have low intracellular levels of multiple lysosomal proteins, accumulate LDL-derived cholesterol (36).

[0160] As stated above, the foregoing is intended to be illustrative of the embodiments of the present invention, and are not intended to limit the invention in any way. Although the invention has been described with respect to the specific modifications described above, the details thereof are not to be construed as limitations, for it will be apparent that various equivalents, changes, and modifications may be resorted to without departing from the spirit and scope thereof and it is understood that such equivalent embodiments are to be included herein.

REFERENCES

1. **Brady, et al.** 1966. The metabolism of sphingomyelin. II. Evidence of an enzymatic deficiency in Niemann-Pick disease. Proc. Natl. Acad. Sci. U.S.A. 55:366-369.

2. **Brady, R.O.** 1978. The Metabolic Basis of Inherited Disease. 4th ed. (Stanbury, J.B., Wyngaarden, J.B., and Fredrickson, D.S., eds.) pp.718-730, McGraw-Hill, New York.

3. **Brady, R.O., Filling-Katz, M.R., Barton, N.W., and Pentchev, P.G.** 1989. Niemann-Pick Disease Types C and D. Neurol. Clin. 7:75-88.

4. **Vanier, M.T., and Suzuki, K.** 1998. Recent advances in elucidating Niemann-Pick C disease. Brain Pathol. 8:163-174.

5. **Pentchev, P.G., Vanier, M.T., Suzuki, K., and Patterson, M.C.** 1995. The Metabolic and Molecular Bases of Inherited Disease, 7th ed. (Scriver, C.R., Beaudet, A.L., Sly, W.S., and Valle, J.B., eds.) pp. 2625-2639, McGraw-Hill, New York.

6. **Christmanou, H., et al.** 2000. Deficient ferritin immunoreactivity in tissues from Niemann-Pick type C patients: extension of findings to fetal tissues, H and L ferritin isoforms, but also one case of the rare Niemann-Pick C2 complementation group. Mol. Genet. Metab. 70:196-202.

7. **Alpy, F., et al.** 2001. The steroidogenic acute regulatory protein homolog MLN64, a late endosomal cholesterol-binding protein. J. Biol. Chem. 276:4261-4269.

8. **Marx, J.** 2001. Disease genes clarify cholesterol trafficking. Science 290:5500.

9. **Vanier, M.T., Duthel, S., Rodriguez-Lafrasse, C., Pentchev, P., and Carstea, E.D.** 1996. Genetic heterogeneity in Niemann-Pick C disease: a study using somatic cell hybridization and linkage analysis. Am. J. Hum. Genet. 58:118-125.

10. **Steinberg, S.J., Ward, C.P., and Fensom, A.H.** 1994. Complementation studies in Niemann-Pick disease type C indicate the existence of a second group. J. Med. Genet. 31:317-320.

11. **Steinberg, S.J., Mondal, D., and Fensom, A.H.** 1996. Co-cultivation of Niemann-Pick disease type C fibroblasts belonging to complementation groups alpha and beta stimulates LDL-derived cholesterol esterification. J. Inherit. Metab. Dis. 19:769-774.

12. **Carstea, E.D., et al.** 1997. Niemann-Pick C1 disease gene: homology to mediators of cholesterol homeostasis. Science 277:228-231.

13. **Davies, J.P., Chen, F.W., and Ioannou, Y.A.** 2000. Transmembrane molecular pump activity of Niemann-Pick C1 protein. Science 290:2295-2298.

14. **Millat, G., et al.** 2001. Niemann-pick disease type c: spectrum of he1 mutations and genotype/phenotype correlations in the npc2 group. Am. J. Hum. Genet. 69:1013-1021.

15. **Kirchhoff, C., Osterhoff, C., Habben, I., and Ivell, R.** 1990. Cloning and analysis of mRNAs expressed specifically in the human epididymis. Int. J. Androl. 13:155-167.

16. **Krull, N., Ivell, R., Osterhoff, C., Kirchhoff, C.** 1993. Region-specific variation of gene expression in the human epididymis as revealed by in situ hybridization with tissue-specific cDNAs. Molec. Reprod. Dev. 34:16-24.

17. **Kirchhoff, C., Osterhoff, C, Young, L.** 1996. Molecular cloning and characterization of HE1, a major secretory protein of the human epididymis. Biol. Reprod. 54:847-856.

18. **Fouchecourt, S., Metayer, S., Locatelli, A., Dacheux, F., and Dacheux, J.L.** 2000. Stallion epididymal fluid proteome: qualitative and quantitative characterization; secretion and dynamic changes of major proteins. Biol. Reprod. 62:1790-1803.

19. **Okamura, N., et al.** 1999. A porcine homolog of the major secretory protein of human epididymis, HE1, specifically binds cholesterol. Biochim. Biophys. Acta 1438:377-387.

20. **Frohlich, O., and Young, L.G.** 1996. Molecular cloning and characterization of EPI-1, the major protein in chimpanzee (Pan troglodytes) cauda epididymal fluid. Biol. Reprod. 54:857-864.

21. **Ellerbrock, K., Pera, I., Hartung, S., and Ivell, R.** 1994. Gene expression in the dog epididymis: a model for human epididymal function. Int. J. Androl. 17:314-323.

22. **Larsen, L.B., Ravn, P., Boisen, A., Berglund, L., and Petersen, T.E.** 1997. Primary structure of EPV20, a secretory glycoprotein containing a previously uncharacterized type of domain. Eur. J. Biochem. 243:437-441.

23. **Nakamura, Y., et al.** 2000. Primary structure, genomic organization and expression of the major secretory protein of murine epididymis, ME1. Gene 251:55-62.

24. **Kirchhoff, C., Osterhoff, C., Pera, I., and Schroter, S.** 1998. Function of human epididymal proteins in sperm maturation. Andrologia 30:225-232.

25. **Naureckiene, S.** 2000. Identification of HE1 as the second gene of Niemann-Pick C disease. Science 290:2298-2301.

26. **Brown, D., and Breton, S.** 2000. H(+)V-ATPase-dependent luminal acidification in the kidney collecting duct and the epididymis/vas deferens: vesicle recycling and transcytotic pathways. J. Exp. Biol. 203:137-145.

27. **Raczek, S., et al.** 2000. Immunocytochemical localisation of some lysosomal hydrolases, their presence in luminal fluid and their directional secretion by human epididymal cells in culture. Cell Tissue Res. 280:415-425.

28. **Tomomasa, H., et al.** 1994. Lysosomal cysteine proteinases in rat epididymis. J. Histochem. Cytochem. 42:417-425.

29. **Marx, J.** 2001. Bad for the heart, bad for the mind?. Science 294:508-509.

30. **Patterson, M.C., et al.** 2001. The Metabolic and Molecular Bases of Inherited Disease, 8th ed. (Scriver, C.R., et al., eds.) pp. 3611-3634, McGraw-Hill, New York.

31. **Sleat, D.E., et al.** 1997. Associations of mutations in a lysosomal protein with classical late-infantile neuronal ceroid lipofuscinosis. Science 277:1802-1805.

32. **Sleat, D.E., Sohar, I., Lackland, H., Majercak, J., and Lobel, P.** 1996. Rat brain contains high levels of mannose-6-phosphorylated glycoproteins including lysosomal enzymes and palmitoyl-protein thioesterase, an enzyme implicated in infantile neuronal lipofuscinosis. J. Biol. Chem. 271:19191-19198.

33. **Kornfeld, S., and Sly, W.S.** 2001. The Metabolic and Molecular Bases of Inherited Disease, 8th ed. (Scriver, C.R., et al., eds.) pp. 3469-3482, McGraw-Hill, New York.

34. **Valenzano, K.J., Kallay, L.M., and Lobel, P.** 1993. An assay to detect glycoproteins that contain mannose 6-phosphate. Anal. Biochem. 209:156-162.

35. **Wattiaux, M., Wabo, P., and Baudhuin, P.** 1963. Ciba Foundation Symposium "Lysosomes". pp. 176-200, Brown, Boston.

36. **Vanier, M.T., et al.** 1991. Type C Niemann-Pick disease: spectrum of phenotypic variation in disruption of intracellular LDL-derived cholesterol processing. Biochim. Biophys. Acta 1096:328-337.

37. **Higgins, M.E., Davies, J.P., Chen, F.W., and Ioannou, Y.A.** 1999. Niemann-Pick C1 is a late endosome-resident protein that transiently associates with lysosomes and the trans-Golgi network. Mol. Genet. Metab. 68:1-13.

38. **Neufeld, E.B., et al.** 1999. The Niemann-Pick C1 protein resides in a vesicular compartment linked to retrograde transport of multiple lysosomal cargo. J. Biol. Chem. 274:9627-9635.

39. **de Duve, C., et al.** 1955. Tissue fractionation studies 6. Intracellular distribution patterns of enzymes in rat liver tissue. Biochem. J. 60:604-612.

40. **Sohar, I., Lin, L, and Lobel, P.** 2000. Enzyme-based diagnosis of classical late infantile neuronal ceroid lipofuscinosis: comparison of tripeptidyl peptidase I and pepstatin-insensitive protease assays. Clin. Chem. 46:1005-1008.

41. **Jadot, M., Wattiaux-De Coninck, R., and Wattiaux, R.** 1985. Effect on lysosomes of invertase endocytosed by rat-liver. Eur. J. Biochem. 151:485-488.

42. **Nishiyama, C., Yuuki, T., Takai, T., Okumura, Y., and Okudaira, H.** 1993. Determination of three disulfide bonds in a major house dust mite allergen, Der f II. Int. Arch. Allergy Immunol. 101:159-166.

43. **Xiang, X., Han, G., Winkelmann, D.A., Zuo, W., and Morris, N.R.** 2000. Dynamics of cytoplasmic dynein in living cells and the effect of a mutation in the dynactin complex actin-related protein Arp1. Curr. Biol. 10:603-106.

What is claimed is:

1. A method of diagnosing Niemann-Pick C2 disease in a subject, comprising detecting at least one mutation in the human HE1 gene sequence of SEQ ID NO: 1.

2. The method of claim 1 wherein said mutation results in at least one mutation of an amino acid in the human HE1 polypeptide sequence of SEQ ID NO: 2.

3. The method of claim 1 wherein said mutation is present in at least one of two copies of said HE1 gene sequence, wherein said two copies are located on two separate chromosomes.

4. The method of claim 3 wherein said mutation indicates the ability of said subject to genetically transmit Niemann-Pick C2 disease.

5. The method of claim 3 wherein said mutation is present in both copies of said HE1 gene sequence.

6. The method of claim 5 wherein said mutation indicates the Niemann-C2 disease in said subject.

7. The method of claim 5 wherein said mutation is homozygous.

8. The method of claim 5 wherein said mutation is heterozygous.

9. The method of claim 1 wherein said mutation is selected from the group consisting of: deletion, substitution, splice mutation, truncation, frame-shift and insertion.

10. The method of claim 1 wherein said mutation is detected by the steps comprising:

(a) nucleic acid amplification; and

(b) subsequent sequence analysis of the nucleic acid.

11. The method of claim 10 wherein said amplification consist of PCR amplification.

12. The method of claim 11 wherein the nucleic acid is a human HE1 gene.

13. The method of claim 11 wherein said mutation is detected by hybridizing:

(a) the HE1 gene of said subject; and

(b) a labeled polynucleotide probe complementary to a portion of the wild type HE1 nucleotide sequence of SEQ ID NO: 1.

14. The method of claim 13 wherein said polynucleotide probe is DNA.

15. The method of claim 13 wherein said label is selected from the group consisting of: a radioactive isotope, biotin, enzyme reporter group, chemiluminescent label, fluorescent label, and an antibody.

16. The method of claim 11 wherein said HE1 gene of said subject is amplified by method selected from the group consisting of PCR and RT-PCR.

17. The method of claim 13 wherein the medium for said hybridization is selected from the group consisting of solid support, liquid solution, in situ, and Southern blot.

18. The method of claim 13 wherein said subject is a fetus.

19. A composition comprising a biologically effective amount of an HE1 polynucleotide.

20. The composition of claim 19 further comprising an acceptable biological carrier.

21. The composition of claim 19 wherein said HE1 polynucleotide sequence is the sequence of SEQ ID NO: 1.

22. The composition of claim 19 wherein said HE1 polynucleotide sequence has at least 70% identity to SEQ. ID. NO: 1, said identity being calculated over the entire length of SEQ. ID. NO: 1.

23. The composition of claim 19 further comprising a vector, which comprises a polynucleotide, wherein said polynucleotide encodes an HE1 polypeptide.

24. The composition of claim 23 further comprising an expression system, which comprises:

(a) a cassette with an HE1 gene,

(b) a vector; and

(c) a biologically acceptable carrier.

25. The expression system of claim 24 wherein said expression cassette contains an HE1 nucleic acid element selected from the group consisting of HE1 DNA, HE1 cDNA, HE1 antisense sequence, and HE1 RNA.

26. The expression system of claim 25 wherein said expression cassette contains one or more elements selected from the group consisting of the following:

(a) a promoter sequence;

(b) a transcription initiation sequence;

(c) one or more enhancer sequences;

(d) one or more selectable elements; and

(e) a reporter gene.

27. The expression system of claim 24 wherein said vector is selected from the group consisting of: a plasmid, cosmid, lambda phage, M13 phage, a retrovirus, lentivirus, adenovirus, herpes simplex virus (HSV), cytomegalovirus (CMV), adeno-associated virus (AAV), papillomavirus, and simian virus (SV40).

28. The expression system of claim 24 wherein said expression system is capable of expressing an HE1 polypeptide substantially identical to SEQ ID NO: 2 in a host cell.

29. A composition comprising a biologically effective amount of an HE1 polypeptide and an acceptable carrier.

30. The composition of claim 29 wherein said HE1 polypeptide sequence is the sequence of SEQ ID NO: 2.

31. The composition of claim 29 wherein said HE1 polypeptide sequence is substantially similar to SEQ. ID. NO: 2.

32. A method of altering intracellular cholesterol levels comprising administering an effective amount of an HE1 therapeutic.

33. The composition of claim 32 further comprising an acceptable biological carrier.

34. The method of claim 32 wherein said HE1 therapeutic is selected from the group consisting of: a HE1 DNA polynucleotide sequence, a HE1 RNA polynucleotide sequence, a HE1 antisense polynucleotide sequence, a HE1 polypeptide and a HE1 protein.

35. The method of claim 32 wherein said levels are decreased by administration of a HE1 therapeutic selected from the group consisting of: a HE1 DNA polynucleotide sequence, a HE1 RNA polynucleotide sequence, a HE1 polypeptide and a HE1 protein.

36. The method of claim 35 wherein said HE1 therapeutic comprises an HE1 polynucleotide sequence identical to SEQ ID NO: 1.

37. The method of claim 32 wherein said HE1 therapeutic comprises an HE1 polynucleotide antisense sequence complementary to SEQ ID NO: 1.

38. The method of claim 37 wherein cholesterol levels are increased by said administration of said HE1 polynucleotide antisense sequence.

39. The method of claim 35, wherein said HE1 therapeutic comprises an HE1 protein sequence substantially similar to SEQ ID NO: 2.

40. The method of claim 32, wherein said administration comprises introducing into a target cell an effective amount of an HE1 polynucleotide sequence.

41. The method of claim 40 wherein said introduction is performed by transfection.

42. The method of claim 41 wherein said introduction is carried out by a method selected from the group consisting of electroporation, microinjection, ballistic introduction or infection, transvection, calcium phosphate transfection, DAEA-dextran mediated transfection, scrape loading, use of a gene gun, cationic lipid-mediated transfection, and lipofection.

43. The method of claim 41 wherein said transfection occurs in vivo.

44. The method of claim 41 wherein said transfection occurs in vitro.

45. The method of claim 41 wherein said transfection takes place as part of an ex vivo procedure.

46. The method of claim 32 wherein said administration is to a human.

47. The method of claim 32 wherein said subject is a fetus.

48. The method of claims 46 and 47 wherein said administration comprises delivering said HE1 therapeutic by a method selected from the group consisting of: intravascular injection, intramuscular injection, and oral administration.

49. The method of claim 32 wherein the alteration of intracellular cholesterol level is used to treat at least one disease selected from the group consisting of Niemann-Pick C2 disease, cardiovascular disease, atherosclerosis, Alzheimer's disease, and diabetes.

50. The method of claim 39 wherein said administration of HE1 protein downregulates endogenous cholesterol synthesis.

51. A method of diagnosing Niemann-Pick C1 disease in a subject, comprising detecting elevated expression levels of the HE1 gene in a cell.

52. The method of claim 51 wherein said elevated expression levels of HE1 are detected by the steps comprising:

(a) measuring the amount of HE1 mRNA in the cells of said subject; and

(b) comparing the results with the amount of HE1 mRNA in the cells of a healthy subject.

53. The method of claim 52 wherein the method of measuring said amount of HE1 mRNA is selected from the group consisting of Northern blot, RT-PCR, and ribonuclease protection assay.

54. The method of claim 51 wherein said elevated expression levels of HE1 are detected by the steps comprising:

(a) measuring the amount of HE1 protein in the cells of said subject; and

(b) comparing the results with the amount of HE1 protein in the cells of a healthy subject.

55. The method of claim 54 wherein the method of measuring said amount of HE1 protein is Western blotting.

56. The method of claim 51 wherein said subject is a human.

57. The method of claim 51 wherein the cell is a fibroblast.

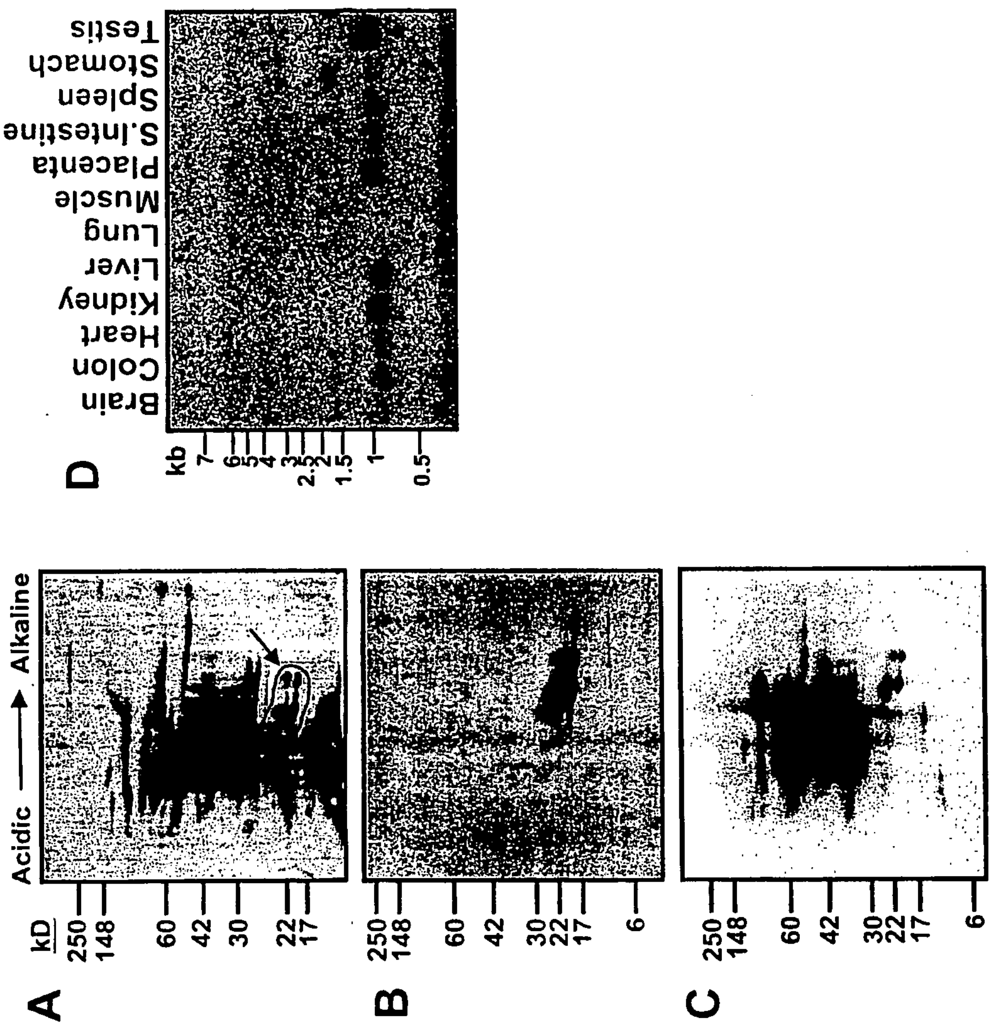
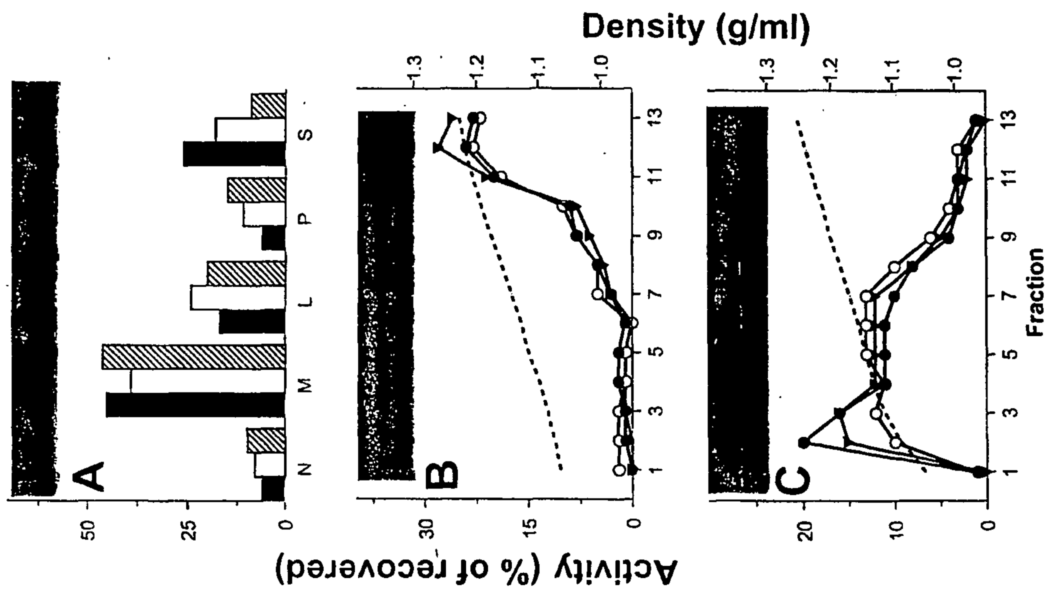

A
B
C
D
kD
250
148
60
42
30
22
17
6
Acidic
Alkaline
kb
7
6
5
4
3
2.5
2
1.5
1
0.5
Brain
Colon
Heart
Kidney
Liver
Lung
Muscle
Placenta
S.Intestine
Spleen
Stomach
Testis

Fig. 1

Fig. 2



BEST AVAILABLE COPY

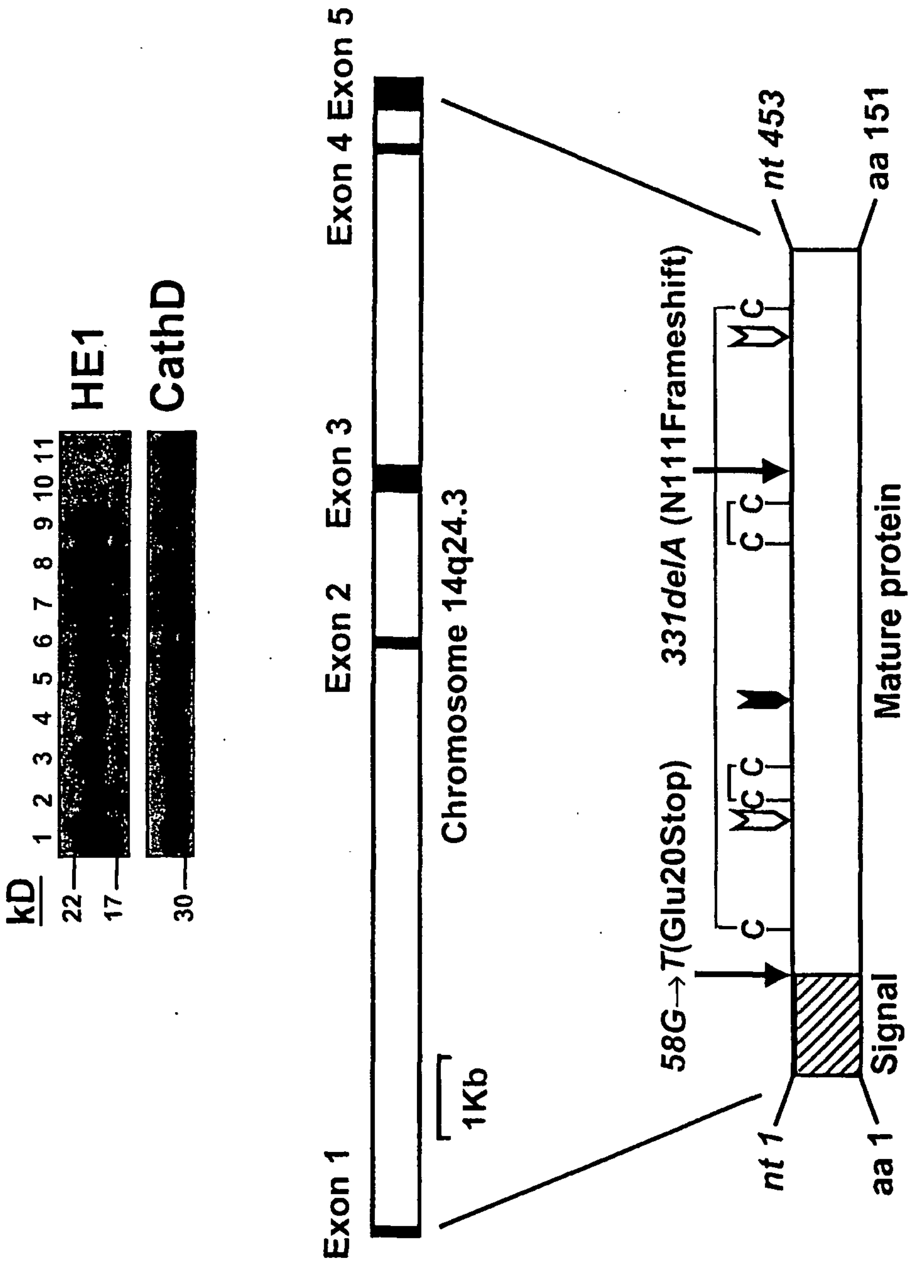

kD
22
17
30
1 2 3 4 5 6 7 8 9 10 11
HE1
CathD
Exon 1
Exon 2
Exon 3
Exon 4 Exon 5
1Kb
Chromosome 14q24.3
58G→T(Glu20Stop)
331delA (N111Frameshift)
C
C C
C C
C
nt 1
nt 453
aa 1
aa 151
Signal peptide (19 aa)
Mature protein (132 aa)

Fig. 3

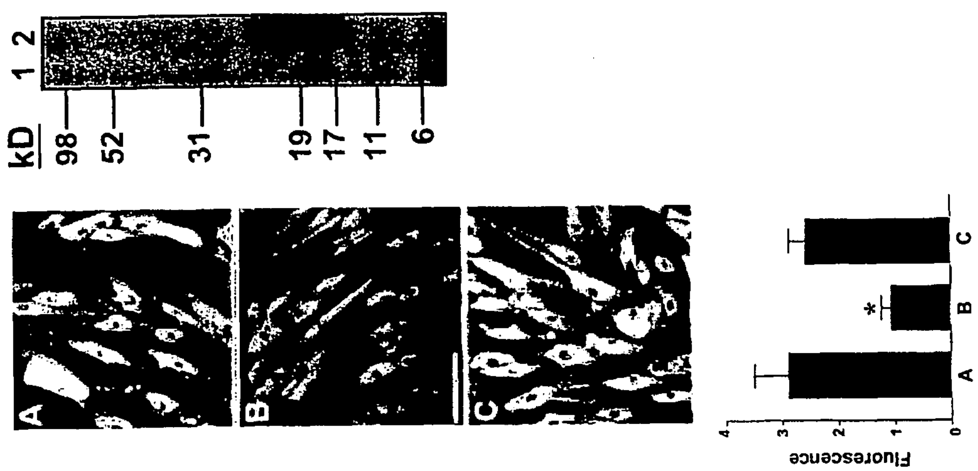


Fig. 4

SEQUENCE LISTING

<110> University of Medical and Dental New Jersey
Lobel, Peter

<120> Compositions and Methods for the Treatment of Diseases Related to Faulty Cholesterol Regulation

<130> 268/211 PCT

<140> 60/245,031
<141> 2000-11-01

<160> 2

<170> PatentIn version 3.1

<210> 1
<211> 808
<212> DNA
<213> Homo sapiens

<220>
<221> gene
<222> (1)..(808)
<223> HE1

<220>
<221> variation
<222> (209)..(209)
<223> allele="T", allele="C"

<220>
<221> variation
<222> (267)..(267)
<223> allele="C", allele="T"

<220>
<221> variation
<222> (688)..(688)
<223> allele="A", allele="C"

<220>
<221> misc_feature
<222> (62)..(451)
<223> note="E1_DerP2_DerF2; Region: E1 family"

<220>
<221> CDS
<222> (11)..(466)
<223>

```
<400>  1
cggattccgg atg cgt ttc ctg gca gct aca ttc ctg ctc ctg gcg ctc        49
           Met Arg Phe Leu Ala Ala Thr Phe Leu Leu Leu Ala Leu
           1               5                   10

agc acc gct gcc cag gcc gaa ccg gtg cag ttc aag gac tgc ggt tct       97
Ser Thr Ala Ala Gln Ala Glu Pro Val Gln Phe Lys Asp Cys Gly Ser
    15                  20                  25

gtg gat gga gtt ata aag gaa gtg aat gtg agc cca tgc ccc acc caa      145
Val Asp Gly Val Ile Lys Glu Val Asn Val Ser Pro Cys Pro Thr Gln
30                  35                  40                  45

ccc tgc cag ctg agc aaa gga cag tct tac agc gtc aat gtc acc ttc      193
Pro Cys Gln Leu Ser Lys Gly Gln Ser Tyr Ser Val Asn Val Thr Phe
                50                  55                  60

acc agc aat att cag tct aaa agc agc aag gcc gtg gtg cat ggc atc      241
Thr Ser Asn Ile Gln Ser Lys Ser Ser Lys Ala Val Val His Gly Ile
            65                  70                  75

ctg atg ggc gtc cca gtt ccc ttt ccc att cct gag cct gat ggt tgt      289
Leu Met Gly Val Pro Val Pro Phe Pro Ile Pro Glu Pro Asp Gly Cys
        80                  85                  90

aag agt gga att aac tgc cct atc caa aaa gac aag acc tat agc tac      337
Lys Ser Gly Ile Asn Cys Pro Ile Gln Lys Asp Lys Thr Tyr Ser Tyr
    95                  100                 105

ctg aat aaa cta cca gtg aaa agc gaa tat ccc tct ata aaa ctg gtg      385
Leu Asn Lys Leu Pro Val Lys Ser Glu Tyr Pro Ser Ile Lys Leu Val
110                 115                 120                 125

gtg gag tgg caa ctt cag gat gac aaa aac caa agt ctc ttc tgc tgg      433
Val Glu Trp Gln Leu Gln Asp Asp Lys Asn Gln Ser Leu Phe Cys Trp
                130                 135                 140

gaa atc cca gta cag atc gtt tct cat ctc taa gtgcctcatt gagttcggtg    486
Glu Ile Pro Val Gln Ile Val Ser His Leu
            145                 150

catctggcca atgagtctgc tgagactctt gacagcacct ccagctctgc tgcttcaaca    546

acagtgactt gctctccaat ggtatccagt gattcgttga agaggaggtg ctctgtagca    606

gaaactgagc tccgggtggc tggttctcag tggttgtctc atgtctcttt ttctgtctta    666

ggtggtttca ttaaatgcag cacttggtta gcagatgttt aattttttttt tttaacaaca    726

ttaacttgtg gcctctttct acacctggaa atttactctt gaataaataa aaactcgttt    786

gtcttgtaaa aaaaaaaaaa aa                                              808
```

<210> 2
<211> 151
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (62)..(451)
<223> note="E1_DerP2_DerF2; Region: E1 family"

<400> 2

Met Arg Phe Leu Ala Ala Thr Phe Leu Leu Leu Ala Leu Ser Thr Ala
1 5 10 15

Ala Gln Ala Glu Pro Val Gln Phe Lys Asp Cys Gly Ser Val Asp Gly
20 25 30

Val Ile Lys Glu Val Asn Val Ser Pro Cys Pro Thr Gln Pro Cys Gln
35 40 45

Leu Ser Lys Gly Gln Ser Tyr Ser Val Asn Val Thr Phe Thr Ser Asn
50 55 60

Ile Gln Ser Lys Ser Ser Lys Ala Val Val His Gly Ile Leu Met Gly
65 70 75 80

Val Pro Val Pro Phe Pro Ile Pro Glu Pro Asp Gly Cys Lys Ser Gly
85 90 95

Ile Asn Cys Pro Ile Gln Lys Asp Lys Thr Tyr Ser Tyr Leu Asn Lys
100 105 110

Leu Pro Val Lys Ser Glu Tyr Pro Ser Ile Lys Leu Val Val Glu Trp
115 120 125

Gln Leu Gln Asp Asp Lys Asn Gln Ser Leu Phe Cys Trp Glu Ile Pro
130 135 140

Val Gln Ile Val Ser His Leu
145 150